Novel pre-stack data confinement and selection for magnetotelluric data processing and its application to data of the Eastern Karoo Basin, South Africa

Anna Platz

Univ.-Diss.

zur Erlangung des akademischen Grades

"doctor rerum naturalium"

(Dr. rer. nat.)

in der Wissenschaftsdisziplin "Angewandte Geophysik"

eingereicht an der

Mathematisch-Naturwissenschaftlichen Fakultät

Institut für Erd- und Umweltwissenschaften

der Universität Potsdam

Gutachter:

1. Gutachter: PD Dr. Ute Weckmann, Universität Potsdam, GFZ Potsdam

2. Gutachter: Prof. Dr. Andreas Junge, Goethe-Universität Frankfurt am Main

3. Gutachter: PD Dr. Jörg Koppitz, Universität Potsdam

Tag der Disputation: 18.07.2018

Published online at the
Institutional Repository of the University of Potsdam:
URN urn:nbn:de:kobv:517-opus4-415087
http://nbn-resolving.de/urn:nbn:de:kobv:517-opus4-415087

Eidesstattliche Erklärung

Hiermit versichere ich, dass ich die vorliegende Dissertation selbständig ohne Hilfe Dritter verfasst habe. Andere als die angegebenen Quellen und Hilfsmittel wurden nicht verwendet. Die den benutzten Quellen wörtlich oder dem Sinn nach entnommenen Abschnitte sind als solche kenntlich gemacht. Dies gilt auch für Zeichnungen, bildliche Darstellungen und dergleichen, sowie für Quellen aus dem Internet. Die Dissertation hat in dieser oder ähnlicher Form weder ganz noch in Teilen einer in- oder ausländischen Hochschule zum Zwecke der Promotion vorgelegen.

Teile von Kapitel 2 wurden bei Geophysical Journal International eingereicht.

Potsdam, den 14.03.2018

Abstract

Magnetotellurics (MT) is a geophysical method that is able to image the electrical conductivity structure of the subsurface by recording time series of natural electromagnetic (EM) field variations. During the data processing these time series are divided into small segments and for each segment spectral values are computed which are typically averaged in a statistical manner to obtain MT transfer functions. Unfortunately, the presence of man-made EM noise sources often deteriorates a significant amount of the recorded time series resulting in disturbed transfer functions. Many advanced processing techniques, e.g. robust statistics, pre-stack data selection or remote reference, have been developed to tackle this problem. The first two techniques reduce the amount of outliers and noise in the data whereas the latter approach removes noise by using data from another MT station. However, especially in populated regions the data processing is still quite challenging even with these approaches. In this thesis, I present two novel pre-stack data confinement and selection criteria for the detection of outliers and noise affected data based on (i) a distance measure of each data segment with regard to the entire sample distribution and (ii) the evaluation of the magnetic polarisation direction of all segments. The first criterion is able to remove data points that scatter around the desired MT distribution and furthermore it can, under some circumstances, even reject complete data cluster originating from noise sources. The second criterion eliminates data points caused by a strongly polarised magnetic signal. Both criteria have been successfully applied to many stations with different noise contaminations showing that they can significantly improve the transfer function estimation. The novel criteria were used to evaluate a MT data set from the Eastern Karoo Basin in South Africa. The corresponding field experiment is part of an extensive research programme to collect information of the current e.g. geological setting in this region prior to a potential shale gas exploitation. The aim was to investigate whether a three-dimensional (3D) inversion of the newly measured data fosters a more realistic mapping of physical properties of the target horizon. For this purpose, a comprehensive 3D model was derived by using all available data. In a second step, I analysed parameters of the target horizon, e.g. its conductivity, that are proxies for physical properties such as thermal maturity and porosity.

Kurzfassung

Magnetotellurik (MT) kann die elektrische Leitfähigkeit des Untergrundes abbilden indem Zeitreihen von natürlichen elektromagnetischen (EM) Wechselfeldern gemessen werden. Während der Datenbearbeitung werden die Zeitreihen in Abschnitte unterteilt und für jeden Abschnitt werden Spektren berechnet, welche auf statistische Art gemittelt werden um MT Übertragungsfunktionen zu bestimmen. Unglücklicherweise beeinflusst die Anwesenheit von künstlichen EM Rauschquellen oft eine signifikante Menge der aufgezeichneten Zeitreihen. Dies führt zu gestörten bzw. falschen Übertragungsfunktionen. Mehrere Methoden wurden entwickelt um dieses Problem zu beheben, z.B. robuste Statistik, pre-stack Datenselektion oder das "remote reference" Verfahren. Die ersten beiden Techniken reduzieren den Anteil von Ausreißern und Rauschen in den Daten während das letzte Verfahren Rauschen mit Hilfe von Daten einer zusätzlichen MT Station entfernt. Trotzdem bleibt die Datenbearbeitung vor allem in besiedelten Gebieten selbst mit diesen Methoden schwierig. In dieser Arbeit präsentiere ich zwei neue pre-stack Datenselektionskriterien zur Bestimmung von Ausreißern und verrauschten Datenpunkten basierend auf (i) einem Distanzmaß unter Berücksichtigung der gesamten Datenverteilung und (ii) der Auswertung der magnetischen Polarisierungsrichtung für jeden Abschnitt. Mit Hilfe des ersten Kriteriums können Datenpunkte entfernt werden, die um die eigentliche MT Verteilung streuen. Außerdem kann es unter bestimmten Umständen sogar ganze Datencluster beseitigen, welche von Rauschquellen hervorgerufen werden. Das zweite Kriterium eliminiert Datenpunkte, welche durch ein stark polarisiertes magnetisches Signal verursacht werden. Beide Kriterien wurden erfolgreich auf viele Stationen mit unterschiedlichen Rauschverhalten angewandt. Weiterhin wurden sie genutzt um MT Daten aus dem östlichen Karoo-Becken in Südafrika auszuwerten. Das dazugehörige Feldexperiment ist Teil eines umfangreichen Forschungsprojektes, welches Informationen über die aktuelle geologische Situation in dem Gebiet sammelt, bevor eine mögliche Schiefergasförderung stattfindet. Der Fokus lag darauf zu untersuchen, ob eine 3D Inversion der neu gemessenen Daten eine realistischere Abbildung der physikalischen Eigenschaften des Zielhorizonts fördert. Zu diesem Zweck wurde ein 3D Modell mit Hilfe aller verfügbaren Daten entwickelt. Anschließend habe ich verschiedene Parameter des Zielhorizonts analysiert, welche Rückschlüsse auf die physikalischen Eigenschaften wie thermische Reife und Porosität erlauben.

Acknowledgements

This thesis would not have been possible without the help and support of a number of people, to only some of whom I can give particular mention here.

Most of all, I want to express my sincere gratitude to my PhD supervisor PD Dr. Ute Weckmann, who gave me the opportunity to write my dissertation. I am thankful for the chance to work on such an interesting topic, to visit South Africa for the first time in my life and to present my work at many national and international conferences. She patiently provided the advice and encouragement necessary for me to go through and finally complete my thesis. She has been a supportive supervisor to me and at the same time gave me the freedom to pursue independent work. Thank you for your patience and support!

Furthermore, I would like to thank Dr. Sissy Kütter, Dr. Gerard Munoz, Dr. Kristina Tietze and Reinhard Klose for introducing me to the MT data processing and all associated programmes, as well as sharing their experiences and insights. This provided me the basis for my work.
I am also thankful to all people who shared their data with me. This gave me the great opportunity to test my novel criteria for many different data sets. Especially, I would like to express my gratitude to Jose Cruces for testing the criteria on his entire data set and for all his helpful comments.

During my PhD time, I had the opportunity to organise a MT field experiment in South Africa. The field work in the Karoo Basin would not have been possible without the help of many people. Many thanks to PD Dr. Ute Weckmann, Manfred Schüler, Dr. Naser Meqbel, Dr. Vierra Wegner, Lucien Bezuidenhout, Sarah Brina, Dr. Bastien Linol, Ashton Dingle, Cedric Patzer, Jade Greve and Warren Miller for collecting the MT data and for the interesting weeks in and around Jansenville. Special thanks to Prof. Moctar Doucouré, Barry Morkel, Stefan Rettig, Manfred Schüler and Gregor Willkommen who helped me a lot to prepare this field experiment. Many thanks to all the tribal communities, farmers and rangers of the game reserve who gave us access to their land and supported us. The instruments for this experiment were provided by the Geophysical Instrument Pool Potsdam (GIPP).

I owe my gratitude to Dr. Naser Meqbel for introducing me to 3D MT modelling and for providing his helpful 3D-Grid programme. He always helped me to progress with my modelling studies and had time for several discussions.
I am also thankful to Dr. Kristina Tietze and Cedric Patzer for answering all my questions concerning inversion strategies, forward modelling and inversion theory.
I would like to thank Jade Greve for her patience when explaining the geology of the Karoo Basin to me. She not only provided me with detailed information but also with delicious African rooibos tea. Special gratitude for all the extended afternoons in which we discussed the different modelling results and possible geological interpretations.
I want to express my special thank to Walja Korolevski and Cedric Patzer for all the fruitful conversations. You always had time to discuss my ideas and to push me in the right direction.
I would like to thank Dr. Kristina Tietze and Dr. Sissy Kütter for many advices related to modelling, the use of programmes and hints concerning practical aspects of the doctorate.
I am thankful to PD Dr. Oliver Ritter for his critical reflection which helped me to question and finally improve my results.

Especially, I would like to acknowledge Marcus Bahrke, Stephanie Lehmann, Jade Greve, Cedric Patzer and Dr. Uwe Döbler who provided helpful feedback on earlier versions of this thesis.

I would like to thank the MT working group (MT-AG), consisting of the Geo-Electromagnetics working groups of Free University Berlin (Dr. Heinrich Brasse) and GFZ Potsdam, as well as the Working Group Applied Geophysics at the University of Potsdam for the many opportunities to present and discuss aspects of my work. I would also like to thank all my present and former colleagues in the Geo-Electromagnetics working group at GFZ for all the discussions, comments and the friendly working atmosphere both in the office and in the field.

I would like to thank the GFZ Potsdam for funding my PhD research.

Finally, I want to thank my family for all their support, patience and help during all these years. This thesis is dedicated to Maman and my grandpa Siegfried.

Contents

List of Figures

List of Tables

Nomenclature

The most important abbreviations and symbols are listed below. Vectors and matrices are written in bold letters. All symbols are also explained in the chapter of their first occurrence.

Abbreviations

$1D$	One-dimensional
$2D$	Two-dimensional
$3D$	Three-dimensional
EMERALD	Electro-Magnetic Equipment, Raw-data And Location Database
$AEON$	Africa Earth Observatory Network
BMA	Beattie Magnetic Anomaly
BP	Bandpass
$C - step$	Concentration step
CFB	Cape Fold Belt
ED	Euclidean distance
EM	Electromagnetic
FT	Fourier Transformation
GFZ	German Research Centre for Geosciences
$Imag$	Imaginary part of a complex number
LSQ	Least squares
MCD	Minimum covariance determinant
MD	Mahalanobis distance
MLE	Maximum likelihood estimator

ModEM	Modular ElectroMagnetic
MPD	Magnetic polarisation direction
MT	Magnetotelluric(s)
MVE	Minimum volume ellipsoid
NMU	Nelson Mandela University
NNMB	Namaqua Natal Mobile Belt
Real	Real part of a complex number
RMS	Root mean square
RR	Remote reference
SS	Single site or single station
VTF	Vertical transfer function

Greek Symbols

α_{B}	Polarisation direction (angle) of the magnetic wave field	[°]
α_{i} $(i = x; y; z)$	Smoothing parameters	
δ	Penetration depth	[m]
λ	Regularisation parameter	
μ_0	Permeability of the free space	[$4\pi{\cdot}10^{-7}$ H/m]
Ω	Regularisation term	
ω	Angular frequency ($\omega{=}2\pi f{=}\frac{2\pi}{T}$)	[s^{-1}]
Φ	Penalty function	
ρ	Electrical resistivity	[Ωm]
ρ_{a}	Apparent resistivity	[Ωm]
σ	Electrical conductivity	[S/m]

φ	Phase	[°]

Latin Symbols

$\overline{x}$	Location	
$\mathbf{B}$	Magnetic flux density	[T]
$\mathbf{C}_{\mathrm{d}}$	Covariance matrix of data errors	
$\mathbf{c}_{\mathrm{i}}$ $(i = x; y; z)$	One-dimensional smoothing and scaling operators	
$\mathbf{C}_{\mathrm{m}}$	Model covariance matrix	
$\mathbf{C}_{\mathrm{x}}$	Covariance matrix	
$\mathbf{d}$	Data vector	
$\mathbf{E}$	Electric field	[V/m]
$\mathbf{M}$	Horizontal magnetic inter-station transfer function	
$\mathbf{m}$	Model parameters	
$\mathbf{m}_0$	Model parameters of the prior model	
$\mathbf{T}$	Vertical (magnetic) transfer function	
$\mathbf{X}$	Data matrix	
$\mathbf{Z}$	Impedance tensor	[m/s]
B_{i} $(i = x; y; z)$	Magnetic field components	[T]
e	Euler's number	
E_{i} $(i = x; y)$	Electric field components	[V/m]
E_{k}	Expected value	
F	Forward operator	
M	Number of model parameters	
$M_{\mathrm{i;j}}$ $(i; j = x; y)$	Inter-station transfer function components	

N	Number of data parameters	
n	Number of events	
p	Number of variables	
${r_\mathrm{b}}^2$	Bivariate quadratic coherence	
T	Period	[s]
T_i $(i = x; y)$	Vertical transfer function components	
x_max	Maximum number of events	
Z_ij $(i; j = x; y)$	Impedance tensor components	[m/s]

Introduction

Magnetotellurics (MT) as a passive geophysical method utilises naturally occurring electromagnetic (EM) fields, caused e.g. by the world wide thunderstorm activity or current systems within the ionosphere. This method images the electrical conductivity structure of the subsurface and can be used for a variety of applications to obtain information about lateral and vertical conductivity variations related to resources and geological processes. For this purpose, long time series of natural EM field variations are recorded at many sites. During data processing, these time series are divided into smaller segments and are subsequently Fourier transformed. The obtained spectral values are then averaged in a statistical manner to estimate different Earth response or so-called transfer functions. Unfortunately, the presence of man-made EM noise sources often deteriorates a significant fraction of the recorded time series resulting in disturbed transfer functions.

Many advanced time series processing techniques have been developed through the years to tackle this problem, e.g. robust statistics, remote reference processing and noise removal in time and/or frequency domain. Nowadays robust statistics are commonly applied for transfer function estimation. Most of the robust statistical algorithms rely on data adaptive weighting schemes and aim to decrease the influence of outliers and noise affected data from a majority of well-behaved samples (e.g. Egbert & Booker, 1986; Chave et al., 1987; Ritter et al., 1998). However, especially in populated regions man-made EM noise superimposes natural EM signals so that a majority or at least a significantly large amount of the time series is deteriorated resulting in disturbed transfer functions even with robust stacking. The remote reference method (Goubau et al., 1978; Gamble et al., 1979) is another established and widely used processing technique to improve the transfer function estimation. The remote reference method requires simultaneously recorded EM fields from a distant reference station. Relevant for a successful application is coherent signal, uncorrelated noise and a reliable and sufficiently accurate time basis for synchronisation. The installation and maintenance of at least one appropriate reference station during a field experiment is not always possible as it requires time and human resources.

In contrast to robust stacking algorithms and remote reference processing, noise removal in time and/or frequency domain reduces noise signals or noise affected data prior to the actual stacking process. For noise removal in time domain often various filters are utilised, e.g. notch and delay line filter (Schmucker, 1978; Chen, 2008) or Wiener filter (e.g. Kappler et al., 2010; Kuetter, 2015), which modify the frequency content or truncate the time series. In frequency domain interactive selection algorithms can be applied to decrease the amount of noise affected data. These approaches are mainly based on physical criteria that can be used to eliminate disturbed parts of the time series

(e.g. Travassos & Beamish, 1988; Weckmann et al., 2005). The application of such tools is usually very tedious, time consuming and requires experienced users.

In the framework of this thesis, I developed two novel pre-stack data confinement and selection criteria. Both of them can lead to significant improvements in transfer function estimation and they work in an almost automatic manner. Both criteria were extensively tested for a variety of stations with different noise contaminations as well as for different transfer functions. Limitations and advantages of the novel criteria are explained using MT data from stations located in southern Africa, Venezuela, Germany and Tajikistan.

The first criterion is statistically based and removes events with a large distance to an estimated data centre under consideration of the covariance matrix. For this purpose, the Mahalanobis distance (MD) is used as a confinement criterion. For a reliable distance calculation, both data centre and covariance matrix have to be estimated by a robust approach. Different algorithms were tested and finally a deterministic minimum covariance determinant algorithm was chosen and implemented within the framework of this thesis. Although this algorithm relies largely on already existing algorithms, it was tailored to the requirements of MT data processing. The success of the MD criterion is limited to cases where the majority of all data is well-behaved as it holds for all statistical approaches. However, in practice their success depends on the required fraction of well-behaved data. If the majority of data points originates from noise sources, this statistical criterion will be ineffective or it will fail. In these cases, additional information or input by the user is necessary to ensure that the majority of all data points represents natural EM signal. However, if the majority of all data is well-behaved, the criterion is able to remove data points scattering around the desired distribution as well as data clusters caused by noise sources.

The second criterion is physically based. The magnetic polarisation direction (MPD) criterion removes events caused by a strongly polarised magnetic signal. The basic idea is that the natural magnetic signal is generated by a variety of sources, e.g. solar activity, ionospheric current systems and lightning, and thus the generated magnetic fields should vary in their incidence directions. Therefore, a preferred polarisation direction is not expected for the magnetic field. For detection of significant polarisation directions all data are evaluated in a histogram and compared to an expected value of a uniform distribution. The criterion has been successfully tested for many cases. Only for very complex polarisation pattern the criterion is not able to reject a sufficient amount of contaminated data.

Both novel criteria were used to improve the data quality of a data set from the Eastern Karoo Basin in South Africa. This data set was measured by the German Research Centre for Geosciences (GFZ)

and the African Earth Observatory Network (AEON) in 2014 and is part of an extensive research programme conducted by AEON.

Within the framework of the ongoing search for new energy resources, shale gas has become important as an alternative resource in the last years. In this context, the shale gas bearing potential of the Karoo Basin, with special focus on the black shales of the Whitehill Formation, has aroused the interest of the petroleum industry and scientists. The research programme conducted by AEON has the aim to obtain information of the current subsurface conditions prior to the exploration and exploitation. This baseline study combines various experiments reaching from groundwater studies, structural geology, different geophysical methods, botanical and zoological subjects, to socio-economical applications.

From previous geophysical studies (e.g. van Zijl, 2006; Weckmann et al., 2007a,b) and laboratory measurements on rock samples (Branch et al., 2007), it is known that the potential shale gas bearing target horizon, the Whitehill Formation, is an electrically conductive marker horizon within more resistive lithological formations of the Karoo Basin. Therefore, the MT method is a prime candidate for imaging the target horizon. At the same time, shallower and deeper aquifers in the region can also be identified. As part of the presented work, a MT field experiment was jointly realised by the GFZ and AEON near Jansenville in November 2014. The location was chosen due to the proximity of two existing shallow boreholes and previous seismic and MT studies. In 2005, Weckmann et al. (2007a) measured MT data along a $70\,km$ long profile in this region. One prominent conductivity anomaly in the two-dimensional (2D) inversion models is a continuous, horizontal band of high conductivity at shallow crustal depths that seems to correlate with the Whitehill Formation. Another prominent feature seated in the middle crust is a narrow, subvertical zone of high electrical conductivity below the centre of the Beattie Magnetic Anomaly (BMA), a fossil, continental scale static magnetic anomaly running through the southern Africa tip.

The newly acquired five component MT data were collected at 111 stations along seven profiles in a period range between $10^{-4} - 10^{3}\,s$ in 2014. In order to obtain a reliable and robust three-dimensional (3D) image of the electrical conductivity structure in this region with special focus of the resolution of the potential shale gas bearing Whitehill Formation, a high quality MD data set is required. 3D inversion in particular is an ill-posed inversion problem, so that it is essential to have a dense station grid and undisturbed data covering a broad period range. The latter issue was addressed by my new development in data processing. In order to add more data, different transfer functions were included into the inversion. Since for exploration, the electrical conductivity of the black shales of the Whitehill Formation is needed as a proxy for the viability of the source rock, conductivity variations in models have to be unambiguously ascribed to real material changes. However, 3D inversions strongly depend

on regularisation, station distribution and other parameter settings. Therefore, I conducted extensive parameter studies and tests with synthetic data to assess limitations in resolution and to provide important information for further experiments.

1 Theory of the magnetotelluric method and its data processing

1.1 Introduction to the magnetotelluric method

Magnetotellurics (MT) is a passive electromagnetic (EM) exploration method for imaging the electrical conductivity structure of the subsurface by measuring natural variations of the magnetic and the electric fields at the Earth's surface. The electrical conductivity or its reciprocal, the resistivity, of geological materials covers more than seven decades and is sensitive to small changes in minor constituents of rocks. Its value does not allow to unambiguously identify a specific rock; different lithologies and rock types have a similar range of resistivities. Their actual value depends on several properties such as temperature, porosity and/or the interconnectivity of minor conductive constituents as e.g. fluids, partial melt or highly conductive minerals such as graphite or sulphide.

The MT method was independently introduced by Tikhonov (1950) and Cagniard (1953) and describes a linear relationship between measured electric and magnetic fields in the frequency domain. The measurement setup of this method is explained in Figure 1.1.

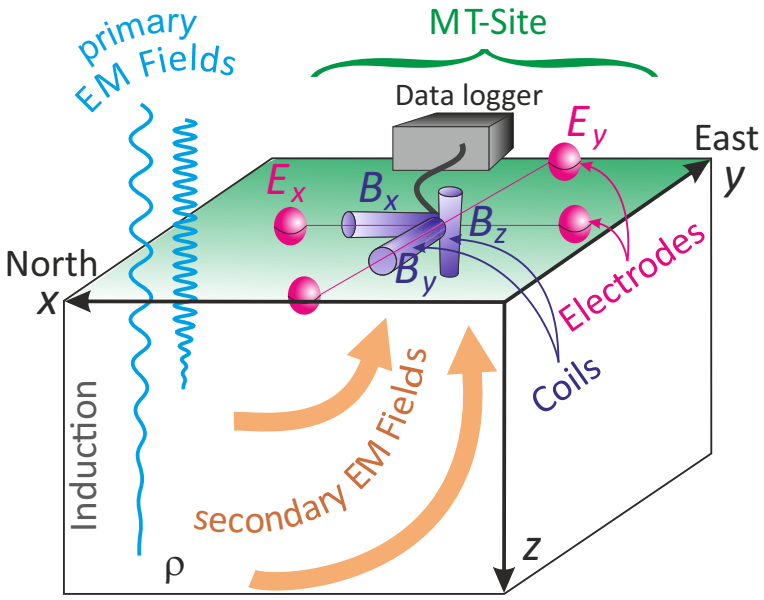

Figure 1.1: Experimental setup of a MT station modified after Sass (2013).

Natural variations of the EM field penetrate into the Earth and induce secondary EM fields. These

secondary EM fields contain information about the electrical conductivity structure of the subsurface. A superposition of primary and secondary EM fields is then measured at a MT station on the Earth's surface. Two orthogonal, horizontal components of the electric field are measured by electrodes and three orthogonal components of the magnetic field are measured by induction coils or fluxgate magnetometers.

The typical period range of naturally generated EM fields in MT applications ranges from $10^{-4}\,s$ to $10^{4}\,s$. Main sources of these EM fields are the global lightning activity (periods $< 1\,s$), current systems and the interaction of the solar wind with the Earth's ionosphere and magnetosphere.

For the period range used in MT and typical electrical conductivities of Earth materials, the propagation of EM energy in the Earth can be described by a diffusion process. The corresponding diffusion equations can be derived from the Maxwell's equations under consideration of simplifying assumptions. Details can be found in a multitude of textbooks, e.g. Kaufmann & Keller (1981); Simpson & Bahr (2005); Chave & Jones (2012); therefore, I refrain from introducing them again.

1.2 Transfer functions

Maxwell's equations imply that the electric and magnetic field components are linked by linear relationships in the frequency domain. Such linear relationships can be described through transfer functions linking an input and an output quantity. Period dependent transfer functions are obtained by MT time series data processing and can subsequently be used to analyse the subsurface electrical conductivity structure.

In this thesis, I used three different transfer functions. The most common MT transfer function is the impedance. It describes the relationship between horizontal electric and horizontal magnetic field components:

$$\begin{pmatrix} E_x \\ E_y \end{pmatrix} = \begin{pmatrix} Z_{xx} & Z_{xy} \\ Z_{yx} & Z_{yy} \end{pmatrix} \begin{pmatrix} B_x \\ B_y \end{pmatrix} \tag{1.1}$$

with $\mathbf{E}$ being the electric field in $[V/m]$, $\mathbf{B}$ the magnetic field in $[T]$ and Z_{ij} $(i; j = x; y)$ the components of the impedance tensor $\mathbf{Z}$ in units of $[m/s]$. The complex numbered 2×2 impedance tensor carries information about the Earth's electrical conductivity structure and each impedance tensor component Z_{ij} can be expressed and visualised as magnitude in terms of apparent resistivity $\rho_{a,ij}$ and phase φ_{ij}:

$$\rho_{a,ij}(\omega) = \frac{\mu_0}{\omega} \left| Z_{ij}(\omega) \right|^2 \tag{1.2}$$

$$\varphi_{ij}(\omega) = \arctan \frac{Imag(Z_{ij}(\omega))}{Real(Z_{ij}(\omega))}. \tag{1.3}$$

The apparent resistivity is the average resistivity of the volume that is penetrated by the EM fields for a given period. The phase value expresses the phase lag between the electric and magnetic fields. Although apparent resistivity and phase do not have tensorial properties any more, these quantities are more descriptive.

While we observe impedances independent of the dimensionality structure of the Earth's subsurface, the vertical magnetic field transfer function emerges only if a lateral conductivity contrast is nearby. The vertical (magnetic) transfer function (VTF) linearly relates this component with the two horizontal magnetic field components:

$$B_z = \begin{pmatrix} T_x & T_y \end{pmatrix} \begin{pmatrix} B_x \\ B_y \end{pmatrix}. \tag{1.4}$$

Typically, this transfer function is graphically represented by induction vectors (or arrows), which are composed by real and imaginary parts of T_x and T_y. In Wiese convention (Wiese, 1962), induction vectors of real parts point away from good conductors. In general, VTFs do not contain information on the absolute values of the subsurface conductivities but they are extremely sensitive to relative conductivity changes. Figures of induction vectors of many sites for one period in map view are common to visualise the presence of lateral conductivity variations in an area.

Similar to the impedance, inter-station transfer functions can always be measured independently of the conductivity structure of the subsurface. This transfer function relates the horizontal magnetic field components of a local station (index l) with the horizontal magnetic fields of a remote station (index r):

$$\begin{pmatrix} B_{x,l} \\ B_{y,l} \end{pmatrix} = \begin{pmatrix} M_{xx} & M_{xy} \\ M_{yx} & M_{yy} \end{pmatrix} \begin{pmatrix} B_{x,r} \\ B_{y,r} \end{pmatrix}. \tag{1.5}$$

There are different options for the remote station: (i) a simultaneously recording station of the experiment layout, (ii) a dedicated reference station or (iii) a synthetic station with e.g. the average horizontal magnetic fields of several MT stations. The inter-station transfer function contains additional information about the conductivity structure of the subsurface and is especially helpful in areas where the subsurface conductivity structure is complex or three-dimensional.

Similar to the inter-station transfer function in equation (1.5) a transfer function $\mathbf{Z}^{'}$ between local

electric and remote magnetic fields can be defined:

$$\begin{pmatrix} E_{x,l} \\ E_{y,l} \end{pmatrix} = \begin{pmatrix} Z'_{xx} & Z'_{xy} \\ Z'_{yx} & Z'_{yy} \end{pmatrix} \begin{pmatrix} B_{x,r} \\ B_{y,r} \end{pmatrix}. \tag{1.6}$$

This transfer function $\mathbf{Z}'$ is sometimes used in pseudo-remote reference processing to replace the normal impedance $\mathbf{Z}$ of a local station, if the local magnetic fields are useless due to e.g. faulty sensors or broken cables. This approach is justified as long as the inter-station transfer function $\mathbf{M}$ between local and remote station is close to the unity matrix.

For different dimensionalities of the subsurface conductivity structure, the transfer functions in equation (1.1), (1.4) and (1.5) show different behaviour (Fig. 1.2).

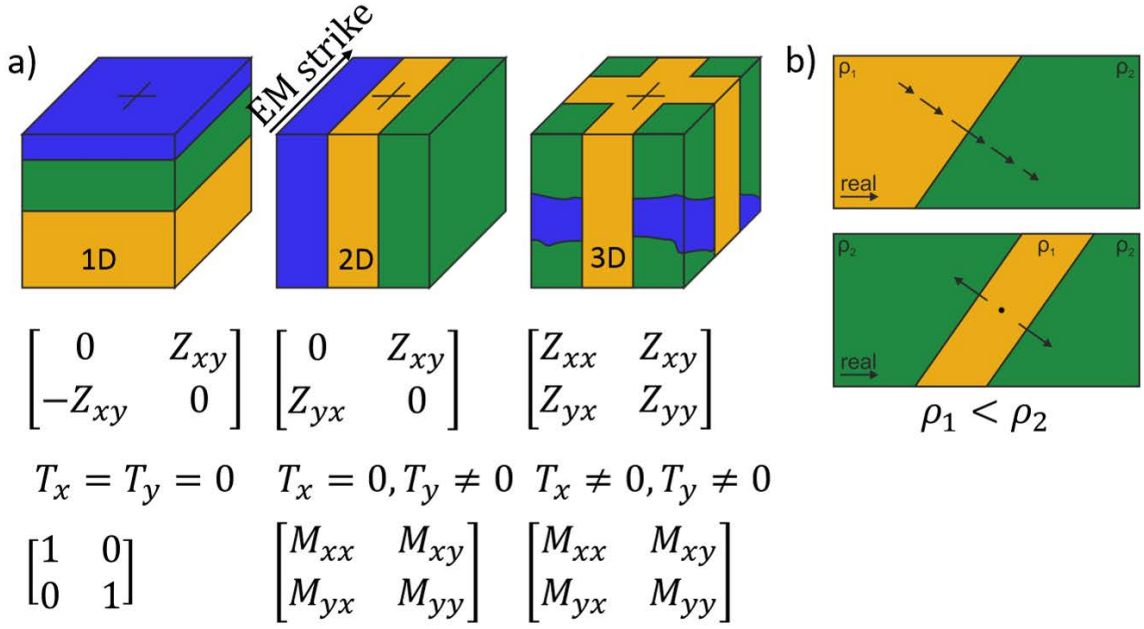

Figure 1.2: a) Behaviour of the different transfer functions for a 1D, 2D and 3D Earth model. The different colours represent zones of different conductivities. b) Expected behaviour for the real induction vectors in Wiese convention.

For a homogeneous half-space, the apparent resistivity represents the true resistivity ρ of the subsurface and the phases are 45° for all periods. For a 1D layered subsurface, both off-diagonal components of the impedance tensor (Z_{xy} and Z_{yx}) have the same value with opposing signs. Impedance phases higher than 45° indicate decreasing resistivity with depth and phases less than 45° are indicative of increasing resistivity. The inter-station transfer function $\mathbf{M}$ is equal to the unity matrix. The VTF does not exist in such an environment. For a 2D subsurface, conductivities vary along one

horizontal direction and with depth. For an arbitrary measurement setup, the impedance tensor is fully occupied. However, it can be mathematically transformed to its simplified form shown in Figure 1.2a, if the measurement's coordinate system is aligned along the EM strike direction.Furthermore, the Maxwell equations can be decoupled into two separate systems of equations (polarisations) called TE- (tangential electric) and TM- (tangential magnetic) mode. The tensor **M** of the inter-station transfer function simplifies depending on the station locations in relation to the conductivity contrast. For the VTF exists only the component perpendicular to the strike direction. For more complex electrical conductivity structures all possible transfer function components exist.

In addition, small-scale conductivity contrasts of the near-surface can significantly distort the impedance tensor. Such localised inhomogeneities cause a galvanic response due to accumulation of charges at the conductivity boundaries. The galvanic distortion of the electric field is constant over period and lead to a quasi-static shift in the apparent resistivity curves. The impedance phase, VTFs and inter-station transfer functions are not affected by this phenomenon. In 2D modelling, this problem is often solved by down-weighting the affected apparent resistivity component, whereas in 3D different approaches have been developed to tackle this problem. In this thesis, I follow the assumption of e.g. Newman et al. (2008) and Xiao et al. (2010), that 3D inversion is capable to deal with static shift effects by integrating compensating structures in the layers near the surface.

All transfer functions contain a certain depth information as the penetration depth of EM fields increases with increasing periods. The penetration depth of EM fields within a homogeneous half-space is defined by its skin depth δ, with

$$\delta = \sqrt{\frac{2}{\mu_0 \sigma \omega}}. \tag{1.7}$$

The skin depth describes the depth at which the amplitude of the initial fields at the surface has decayed to $1/e$ of its original value.

1.3 Introduction to the magnetotelluric data processing

The aim of the MT data processing is to estimate the different transfer functions mentioned in section 1.2 from the recorded time series. In the framework of this thesis, I used the processing package EMERALD (Ritter et al., 1998; Weckmann et al., 2005; Krings, 2007). The main work flow of the EMERALD software package is quite similar to other MT processing algorithms. Time series of typically five component EM field variation data are bandpass filtered and divided into short, contiguous time

windows of fixed lengths. Subsequently, these time windows are Fourier transformed, corrected for the instrument response functions and averaged onto target periods that are for example equally distributed on a logarithmic scale. From smoothed auto- and cross-spectra different transfer functions are estimated. The entire processing scheme is summarised in Figure A.1 in the appendix.

1.4 Estimation of transfer functions

For transfer function estimation, equations (1.1), (1.4), (1.5) and (1.6) have to be solved. All of them can be mathematically treated in the same manner and thus I will use a simplified term:

$$Z = aX + bY. \tag{1.8}$$

This equation has a linear, bivariate structure with a single output channel Z and two independent input channels X and Y. The output channel Z is associated with either E_x, E_y or $B_{x,l}$, $B_{y,l}$ or B_z for the row-wise solution of the impedance, inter-station or vertical transfer function. Furthermore, the input channels X and Y are associated with the horizontal magnetic fields B_x and B_y of the local or of the remote station for the inter-station transfer function. The different transfer function components a and b are estimated from measured and therefore imperfect data due to the finite sample size and the presence of noise. Therefore, a noise term δZ is added to the right-hand side of equation (1.8).

Within the course of solving equation (1.8), typically least squares (LSQ) methods are applied to obtain values that are statistically averaged over the entire time series divided into time windows. The frequency data of each of these windows are treated as individual samples. The aim is to minimise ∂Z for optimal values of a and b by solving the following minimisation problem:

$$min \left\| Z - aX - bY \right\|_2^2 \tag{1.9}$$

with the subscript 2 denoting the Euclidean norm. The solutions for the transfer function components a and b using the LSQ method are:

$$a = \frac{\langle YY^* \rangle \langle ZX^* \rangle - \langle YX^* \rangle \langle ZY^* \rangle}{\langle XX^* \rangle \langle YY^* \rangle - \langle XY^* \rangle \langle YX^* \rangle} \tag{1.10}$$

$$b = \frac{\langle XX^* \rangle \langle ZY^* \rangle - \langle XY^* \rangle \langle ZX^* \rangle}{\langle XX^* \rangle \langle YY^* \rangle - \langle XY^* \rangle \langle YX^* \rangle} \tag{1.11}$$

with $\langle\rangle$ representing the stacked auto- and cross-spectra, e.g. $\langle ZX^* \rangle = \sum_i (Z_i X_i^*)$ for the ordinary

LSQ method or the weighted stacked spectra for the weighted LSQ method with e.g. $\langle ZX^* \rangle = \sum_i w_i \left(Z_i X_i^* \right)$ and the asterisk denotes a complex conjugate.

When solving equation (1.8) with a LSQ approach, noise is only allowed and accommodated in the output channel Z. Furthermore, the LSQ approach is only the maximum likelihood estimator (MLE) if the EM noise is independent and Gaussian distributed. However, we often find that natural EM variations are overprinted by larger man-made EM signals violating our assumptions, leading to large outliers and as a result disturbing the transfer function estimation. Often, simple single station LSQ methods cannot be used to estimate the transfer functions (Swift & Moore, 1967; Sims et al., 1971) as a significant portion of the time segments distorts the calculated mean value.

The remote reference method is an advanced technique for transfer function estimation, explained in more detail in the next section. In the remote reference formulation (for impedance and VTFs) the auto-spectra of the local fields are replaced by cross-spectra between local and remote fields, leading to following equations for a_R and b_R:

$$a_R = \frac{\langle YY_R^* \rangle \langle ZX_R^* \rangle - \langle YX_R^* \rangle \langle ZY_R^* \rangle}{\langle XX_R^* \rangle \langle YY_R^* \rangle - \langle XY_R^* \rangle \langle YX_R^* \rangle} \tag{1.12}$$

$$b_R = \frac{\langle XX_R^* \rangle \langle ZY_R^* \rangle - \langle XY_R{}^* \rangle \langle ZX_R^* \rangle}{\langle XX_R^* \rangle \langle YY_R^* \rangle - \langle XY_R^* \rangle \langle YX_R^* \rangle}. \tag{1.13}$$

1.5 Robust processing techniques

The measured data consist of the true MT signal and noise. There exist a broad variety of EM noise sources, see e.g. Szarka (1988) and Junge (1996). In addition, outliers can exist originating from noise or the true signal. An outlier is usually characterised as a data point that is different from the remaining data, e.g. extreme values. Robust and advanced processing techniques have been developed through the years to tackle the problem that the presence of man-made EM noise sources as well as large outliers often deteriorates a significant fraction of the recorded time series resulting in disturbed transfer functions. Thereby, three main developments have mainly improved MT data processing: (i) robust statistics, (ii) remote reference processing and (iii) noise removal in time and/or frequency domain.

Nowadays, robust statistics are commonly applied for transfer function estimation. Most of the robust statistical algorithms rely on data adaptive weighting schemes and aim to decrease the influence of outliers and noise affected data (e.g. Egbert & Booker, 1986; Chave et al., 1987; Ritter et al., 1998).

The limiting fraction of outliers and noise a robust estimator can handle is defined as its so-called breakdown point that normally cannot exceed 50 % that means that the majority of data has to be well-behaved (Huber, 1981; Hampel, 1986). Many robust algorithms rely on M-estimators (Huber, 1981; Hampel, 1986) and consequently have breakdown points much smaller than 50 % (Smirnov, 2003). Robust data processings with a higher breakdown point were presented by Smirnov (2003) and Chave & Thomson (2004). In MT, robust estimators are almost always based on a statistical model assuming the majority of segments has a Gaussian core of data and only a minor fraction contains noisy data resulting in outliers whose influence is removed through data processing. However, as residuals from robust estimators are often systematically long tailed in comparison to Gaussian distributed ones, the model assumption of conventional robust estimators is not valid and these robust estimators lack the optimality properties of a MLE. Alternatively, Chave (2014) presented a MLE based on alpha stable distributions. This MLE is based on a model directly derived from the data. In contrast, conventional robust estimators use an a-priori statistical model that is inconsistent with actual MT data.

The robust stacking within EMERALD assumes a Gaussian core of data as most of the conventional robust estimators. The robust algorithm consists of an iterative weighting scheme and is based on two successive algorithms. The first algorithm consists of the so-called chi-square criterion and a z-transformation. In this part, it is examined whether a single event spectrum fits into the global view of the majority of all data and accordingly the weight of a single event spectrum is increased or decreased (Ritter et al., 1998). The second algorithm is called consistency criterion. It iteratively replaces a certain amount of bad data with predicted values and therefore reduces non-stationary contributions in the transfer functions (Ritter et al., 1998). Usually the processing is most effective, when both algorithms are used in combination. The robust stacking algorithm is able to deal with extreme outliers and noise in the tail/rim of the distribution. However, especially in populated regions man-made EM noise superimposes natural EM signals so that a majority or at least a significantly large amount of the time series is deteriorated resulting in disturbed transfer functions even with robust stacking.

The remote reference method (Goubau et al., 1978; Gamble et al., 1979) is another well established and widely used processing technique to improve the transfer function estimation. The remote reference method requires simultaneously recorded EM fields from at least one reference station. Relevant for a successful application is coherent signal, uncorrelated noise and a reliable and sufficiently accurate time basis for synchronisation. In contrast to single site processing, remote reference processing utilises an errors-in-variables model and therefore accommodates for noise in all channels (Chave & Jones, 2012). However, it does not specifically allow for noise in the remote channels. Therefore, remote reference

estimators show significant dependence on the amount of this noise (Egbert, 1997). Furthermore, the installation and maintenance of an appropriate reference station during a field experiment is not always possible as it requires time and human resources.

In contrast to robust stacking algorithms and remote reference processing, noise removal in time and/or frequency domain removes noise signals or noise affected data prior to the actual stacking process. For noise removal in time domain often various filters are utilised, e.g. notch and delay line filter (Schmucker, 1978; Chen, 2008) or Wiener filter (e.g. Kappler et al., 2010; Kuetter, 2015), which modify the frequency content or truncate the time series. In frequency domain, interactive selection algorithms are often applied as a pre-stack tool to decrease the amount of noise affected data. These approaches are mainly based on visual inspection of physical criteria that can be used to eliminate disturbed parts of the time series (e.g. Travassos & Beamish, 1988; Weckmann et al., 2005). The application of such tools is usually very tedious, time consuming and requires experienced users.

So far, best processing results are often only obtained by a combination of above mentioned approaches (Larsen, 1989; Jones et al., 1989; Oettinger et al., 2001; Chave & Thomson, 2004; Weckmann et al., 2005). In addition to the above listed improvements, optional pre-stack data selection criteria can be used in the EMERALD package for further data improvement. Often a coherence threshold is applied, i.e. the bivariate quadratic coherence r_b^2 for each single event is computed:

$$r_b^2 = \frac{a\,[XZ^*] + b\,[YZ^*]}{[ZZ^*]}. \tag{1.14}$$

The bivariate quadratic coherence is basically the ratio of predicted to measured signal energy between output and input channel under the assumption of a linear relationship between them and is not supposed to fall below a given limit (Weckmann et al., 2005). The coherence values can lie in the range of $[0, 1]$. Single events with a smaller coherence value than a user defined threshold are rejected from the further processing, due to the fact that noisy data often do not fulfil the relationship resulting in low coherence values. Therefore, the coherence value of a single event is often regarded as a good indicator of data quality. However, often near field EM noise is also very coherent; in such cases a coherence criterion is counterproductive.

Despite all these robust techniques, we still have MT sites and/or period ranges with insufficient data quality. Therefore, additional and novel approaches are necessary to improve the transfer function estimation. In the framework of my thesis, I implemented two novel pre-stack data confinement and selection criteria into EMERALD, which are described in the next chapter.

Chapter summary:

- Transfer functions contain information about the electrical conductivity structure.
- During the data processing, transfer functions are calculated as average values of smoothed auto- and cross-spectra.
- Several techniques were developed to improve the transfer function estimation, e.g. robust statistics, advanced processing techniques and pre-stack tools.

2 New data confinement and selection criteria

An important step within the MT data processing is the calculation of transfer functions as a weighted average value over the entire time series divided into segments or so-called events. Unfortunately, nowadays natural EM variations are often superimposed by larger man-made EM signals. Especially in populated regions, EM noise often deteriorates a majority or at least a significantly large amount of the time series. In these cases, the application of pure robust stacking will result in disturbed transfer functions. Interactive selection algorithms can be used as a pre-stack tool to decrease the amount of noise affected data prior to the actual stacking process. However, as they rely on a visual inspection of different e.g. physical parameters, their application is usually very tedious, time consuming and requires experienced users. Another way to improve the signal-to-noise ratio prior to the stacking procedure is the application of data confinement and selection criteria. In the framework of this thesis, I developed two such criteria: (i) the Mahalanobis distance (MD) criterion as a statistical criterion that removes events with a large distance to an estimated data centre and (ii) the magnetic polarisation direction (MPD) criterion as a physically based criterion that rejects events belonging to a strongly polarised magnetic signal. Both criteria work in an almost automatic manner and can significantly improve the transfer function estimation.

The first criterion is based on a distance measure to detect outliers and to confine the data to an ideally noise-free subset that is subsequently used in the robust stacking algorithm. An outlier or a data point from an EM noise cluster is usually defined by a certain - normally Euclidean - distance to the mean value of the desired MT data distribution. In MT, different quantities can be considered as data for this statistical analysis, i.e. Fourier coefficients, auto- and cross-spectra or transfer function components. The Fourier coefficients and the derived smoothed auto- and cross-spectra include electric and/or magnetic fields that inherent different units and therefore can exhibit a different variability.

The Euclidean distance (ED) has the drawback that it does not account for different metrics of the individual variables or quantities and accordingly the quantity with the largest range will dominate the result. Furthermore, it cannot be excluded that the used MT data do not exhibit an internal correlation. As the ED does not correct for any correlations, it can be distorted if the data are correlated. Moreover, due to the fact that the ED only considers the mean value of a data distribution but does not account for its shape or topology, it is not an appropriate measure of distance for MT data.

Instead of the ED an advanced distance measure, the Mahalanobis distance (MD), is used. The MD is a commonly used distance measure in multivariate statistics. It is often applied for outlier detection in a wide spectrum of fields reaching from biology and chemistry to lean manufacturing (de Maesschalck et al., 2000; Filzmoser et al., 2005; Srinivasaraghavan & Allada, 2006; Malisa, 2010; Brereton, 2015). Here, I use it as a pre-stack data confinement criterion.

In contrast to the ED, the MD allows for different scaling of the individual variables and any correlation between them by using the covariance matrix in addition to the mean value to describe the shape of the data distribution. Unfortunately, outliers have a strong influence on the estimation of the mean value (called location subsequently) and the covariance matrix (also referred to as scatter). Because of this, it is essential for an effective MD calculation that both quantities are estimated in a robust manner. In the past, several methods were developed for robust multivariate location and scatter estimation. The simplest methods are based on median absolute deviation (Gnanadesikan & Kettenring, 1972; Huber, 1981; Falk, 1997; Friebel et al., 2010) and more complex algorithms are e.g. the minimum volume ellipsoid (MVE) or the minimum covariance determinant (MCD) method (Rousseeuw, 1984, 1985). Several of these approaches were tested and the results are presented in the following sections. The finally implemented MD criterion uses a deterministic MCD algorithm for the robust MD calculation.

However, the MD criterion as a purely statistical criterion is limited to cases where the majority of data is well-behaved. The new criterion fails if the majority of data originates from noise sources. In these cases, some noise has to be removed manually e.g. by interactive selection algorithms, other a-priori information or other data selection criteria.

I expanded the classical MD approach by using the polarisation direction of the magnetic signal (MPD criterion). The MPD is one of the parameters in the interactive selection algorithm from Weckmann et al. (2005). The electric field can exhibit preferred polarisation directions because of a given conductivity distribution of the subsurface. However, a preferred polarisation direction is not expected for the magnetic field as it is generated by a variety of different sources. This fact is used to develop an automatic selection criterion, which removes events that originate from strongly polarised magnetic fields. Thus, the criterion is able to reject events originating from coherent noise sources. In combination with the MD criterion, it can lead to significantly improved processing results for stations that are highly affected by noise.

Both criteria and their implementation are described in the following sections of this chapter. Furthermore, processing results from MT stations located in South Africa, Namibia (Kapinos et al., 2016),

Tajikistan (Korolevski et al., 2014), Germany (Muñoz et al., 2010) and Venezuela (Schmitz et al., 2013) are presented. These results are used to emphasise advances and limitations of the new criteria. The parts of this chapter related to the MD criterion have been submitted by Platz & Weckmann (In revision).

2.1 Introduction of the Mahalanobis distance (MD)

The MD, introduced by the Indian mathematician P.C. Mahalanobis in 1936 (Mahalanobis, 1936), is a commonly used distance measure to detect outliers and defines the distance between a specific multivariate data point and the location of the data set considering the correlation of these data. Mathematically, this definition is given by the following equation:

$$MD_i = \sqrt{(\mathbf{x}_i - \overline{\mathbf{x}})\mathbf{C}_x^{-1}(\mathbf{x}_i - \overline{\mathbf{x}})^T} \tag{2.1}$$

with $\mathbf{x}_i$ being a row vector of the $i-th$ observation or event, $\overline{\mathbf{x}}$ as the location of the sample distribution, e.g. estimated by mean or median, and $\mathbf{C}_x^{-1}$ as the inverse of its covariance matrix. All data can be summarised in a $n \times p$ matrix $\mathbf{X}$ with n as the number of observations or events and p the number of variables or quantities. The MD describes the distance between a multivariate point $\mathbf{x}_i$ and the location of the distribution in terms of multiples of standard deviations. The MD is unitless and affine invariant.

The difference between the MD and the classical ED is demonstrated in Figure 2.1.

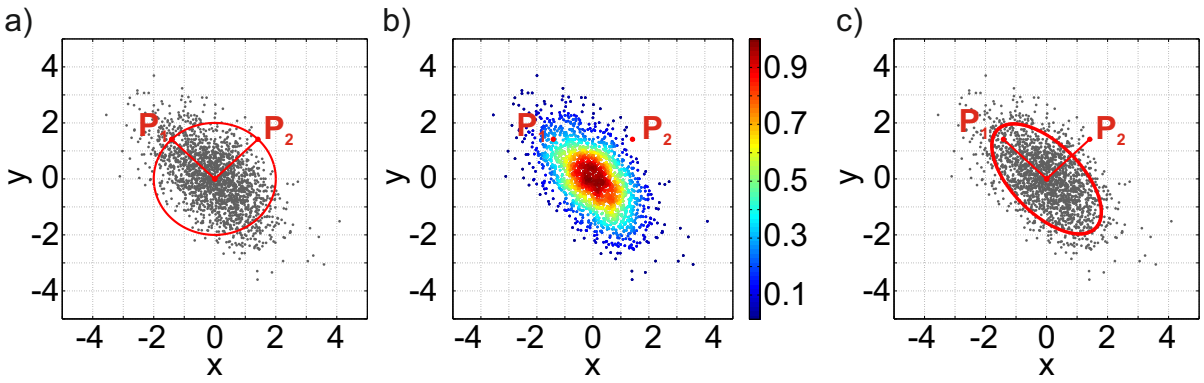

Figure 2.1: Synthetic example of bivariate normally distributed data after Lohninger (2012). a) Points with the same ED are located on the red circle, e.g. points P_1 and P_2. b) The same data distribution, where each data point is colour-coded with its likelihood of occurrence (red $\hat{=}$ high, blue $\hat{=}$ low likelihood). c) Points with the same MD are located on the red ellipse. This example illustrates that the MD better describes correlated data.

For this synthetic example 2000 data points were generated following a bivariate normal distribution. Usually, a bivariate normal distribution is characterised by its mean vector and its covariance matrix, which includes information of the standard deviations for each variable or spatial direction and any correlation between the different components. The mean vector of this data set was set to $\mathbf{0}$ and the standard deviations of the $x-$ and $y-$direction were set to 1 and 1.2, respectively. A correlation coefficient of -0.6 was applied in order to generate linearly correlated data.

Assessing the distribution by the ED, points with the same distance from the data centre are located on a circle, such as the two exemplary points P_1 and P_2 in Figure 2.1a. Although both points have the same ED, the intuitive perception tells us that the occurrence of point P_1 is much more likely than the occurrence of point P_2. To visualise the likelihood of occurrence, I refer to an approach by Eilers & Goeman (2004), where each data point is colour-coded with its smoothed likelihood of occurrence. It represents the empirical distribution of points instead of plotting individual points, since scatterplot presentations with a large number of data points do not allow to assess the actual amount of data points in the central part. The representation in terms of a two-dimensional histogram often results in choppy figures or wide bins. In Figure 2.1b, points with a high likelihood are marked in red, whereas points with a low likelihood are presented in blue colours. It indicates that the likelihoods of the two points differ and that the occurrence of point P_1 is more likely than the occurrence of point P_2 confirming the intuitive perception. Data points with the same MD are located on an ellipse or an ellipsoid in higher dimensions as shown in Figure 2.1c. In comparison to the red circle of the ED, the red ellipse better summarises points with a similar likelihood of occurrence and therefore the MD is a better distance measure.

For uncorrelated data, which have the same standard deviations in each spatial direction or for each variable, the MD simplifies to the ED. In MT data processing, Fourier coefficients and/or auto- and cross-spectra are used. These quantities can span different ranges as they are derived from electric and/or magnetic fields, which have different units. Furthermore, the real and imaginary parts of these quantities or of the different transfer function components can have different standard deviations. Moreover, internal correlations between the used data cannot be excluded. Therefore, instead of the ED, the MD is used to detect and remove outliers as well as to confine the data to a subset that is subsequently used in the stacking process.

2.2 Data confinement with the Mahalanobis distance

The real and imaginary parts of the transfer function components in equation (1.8) are used as variables for the MD analysis. Weckmann et al. (2005) showed that solving for each field component individually, e.g. each electric field component for impedance, often results in smoother transfer functions in case of EM noise affecting only one component. Such a strategy results in $p = 4$ variables for the data matrix $\mathbf{X}$. Within the EMERALD processing, equation (1.8) is solved for each target period individually. Consequently, the number of observations n is given by the number of events for the examined period. For each single event, a MD value can be calculated by using equation (2.1). As mentioned in the introduction of this chapter, outliers as well as data points from EM noise clusters are usually defined by a certain distance to the location of the desired MT data distribution. Therefore, events with a large MD value are considered as potential outliers or noise affected data and are removed from the further processing.

The definition of a maximal allowed MD threshold is essential to confine the data in an automatic manner. There exist several ways to define such a threshold. The best way is to define a threshold by visual inspection of the distribution of all MD values for a given period. However, this approach is very time consuming and cannot be realised in an automatic manner. Therefore, other approaches are often used to estimate an appropriate threshold.

In most cases, the threshold is defined by a certain quantile of the χ^2-distribution. This idea is based on the fact that the squared MDs are approximately χ^2-distributed if the data follow a multivariate normal distribution. The degrees of freedom of the χ^2-distribution are determined by the dimension p of the multivariate data.

The $\alpha-$quantile of a probability distribution P is the value x_α, which divides the distribution into two intervals so that

$$P((-\infty, x_\alpha]) \geq \alpha \;\&\; P([x_\alpha, +\infty)) \geq 1 - \alpha \tag{2.2}$$

holds with $\alpha \in (0, 1)$. Typical values for $\alpha-$quantiles and derived MD thresholds are listed in Table 2.1 for a χ^2-distribution with four degrees of freedom.

α	0.500	0.750	0.900	0.950	0.975	0.990	0.995
x_α	3.36	5.39	7.78	9.49	11.14	13.28	14.86
MD threshold	1.8	2.3	2.8	3.1	3.3	3.6	3.9

Table 2.1: $\alpha-$quantiles for a χ^2-distribution with four degrees of freedom from Morrison (1967) and the derived MD thresholds (root of x_α) for standard EMERALD processing. The MD thresholds are rounded to the first decimal place.

If MT data are noise-free, i.e. containing no outliers, then the squared MD values lie with a probability α in the interval $[0, x_\alpha]$. Therefore, events with a MD value greater than the selected threshold are identified as potential outliers originating e.g. from noise sources, and these events are removed from the further processing.

An alternative approach to define a threshold without making assumptions about the data distribution is given by the Chebyshev's inequality also known as Chebyshev theorem. This approach has the advantage that it allows the explicit calculation of thresholds without knowing the data distribution. The Chebyshev theorem, formulated by the Russian mathematician P.L. Tschebyscheff, can be used to estimate a lower bound of the percentage P of data that lie within k standard deviations σ_s from the mean μ (Amidan et al., 2005; Lohninger, 2012):

$$P\left(|x-\mu| \leq k\sigma_s\right) \geq \left(1-\frac{1}{k^2}\right). \tag{2.3}$$

From this equation follows that e.g. with $k = 3.9$ at least 93.43 % of the data would fall within 3.9 standard deviations from the mean for univariate data.

In practice, this way to define thresholds is not often used because it requires the standard deviation σ_s and the mean value μ, which are normally not known a-priori and which have to be estimated from the data itself. Using the arithmetic mean and the sample standard deviation as estimates for these two quantities in equation (2.3) can lead to unreliable results due to the fact that both quantities are sensitive to outliers. Therefore, robust estimates are essential for the estimation of a reliable threshold.

There exist several extensions of equation (2.3) to the multivariate case, see e.g. Marshall & Olkin (1960); Chen (2007); Stellato et al. (2016). The extension proposed by Chen (2007) is related to the MD:

$$P\left\{\left(X-E\left[X\right]\right)^T \Sigma^{-1}\left(X-E\left[X\right]\right) \geq \epsilon\right\} \leq \frac{p}{\epsilon}, \quad \forall \epsilon > 0 \tag{2.4}$$

with $(X-E\left[X\right])^T \Sigma^{-1}(X-E\left[X\right])$ being the squared MD. The confidence intervals derived from this equation are ellipsoids, which are centred around the population mean $E\left[X\right]$. These ellipsoids represent points with the same MD.

Due to the fact that the Emerald processing assumes an underlying Gaussian model, I calculated thresholds by using $\alpha-$quantiles of a χ^2-distribution. The corresponding threshold defined by the Chebyshev theorem would be always larger, because of the independence of an assumed distribution. Although it is known that MT data are rather stably than normally distributed (Chave, 2014), tests

with different stations and data qualities suggest that the thresholds derived from $\alpha-$quantiles of the χ^2-distribution are adequate for MT data. Therefore, I used the values of table 2.1 as orientation for defining the thresholds in the following examples.

2.3 Estimation scheme of the Mahalanobis distance criterion

Based on the idea that the MD is an appropriate measure to identify outliers, I used the MD to develop a pre-stack data confinement criterion. This novel MD criterion was implemented into the processing package EMERALD (Ritter et al., 1998; Weckmann et al., 2005; Krings, 2007). MT transfer functions are thereby estimated from smoothed auto- and cross-spectra for each target period. This estimation is done in a robust manner by using the iterative robust weighting scheme after Ritter et al. (1998).

Additional pre-stack selection criteria, such as the physically motivated coherence criterion, can be used in EMERALD. In the following, I refer to the original EMERALD processing with the term "standard EMERALD processing", while the usage of the novel MD criterion is subsumed under "EMERALD+MD processing".

The MD criterion is implemented after incoherent data are removed and before the robust stacking algorithm removes outliers. In contrast to physically motivated data selection criteria, the MD criterion is based on statistics. Main steps and the kernel routine of this new criterion are explained in the next two subsections and are summarised in Figure A.2 in the appendix.

2.3.1 Robust calculation of Mahalanobis distances

For a reliable outlier and noise detection using the MD in equation (2.1), two prerequisites have to be fulfilled: (i) the calculation of the covariance matrix requires more samples or events than variables (de Maesschalck et al., 2000; Brereton, 2015), this means that at least five events per period are needed to use real and imaginary parts of the complex numbered transfer functions (see eq. 1.8) as variables for the MD computation. (ii) A robust calculation of the MD is vital.

The computation of the MD depends on the location $\overline{\mathbf{x}}$ and the covariance matrix $\mathbf{C}_{\mathrm{x}}$, which have

to be estimated from the data set and therefore strongly depend on outliers within these data. A straightforward calculation of these two quantities could be realised by the arithmetic mean and the sample or empirical covariance matrix:

$$\overline{\mathbf{x}} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{2.5}$$

$$\mathbf{C}_x = \frac{1}{n-1}\sum_{i=1}^{n}(x_{ij} - \overline{\mathbf{x}_j})(x_{ik} - \overline{\mathbf{x}_k}) \tag{2.6}$$

with $j, k = 1, \ldots, p$ representing different variables. Unfortunately, this approach is not robust and MDs calculated by these two quantities can be highly misleading due to the large sensitivity of the arithmetic mean and the covariance matrix to outliers (Rousseeuw & van Driessen, 1999; Filzmoser et al., 2005; Hubert & Debruyne, 2010; Lehmann, 2012).

Many robust algorithms for robust location and scatter estimation can be found in literature, see e.g. Gnanadesikan & Kettenring (1972); Huber (1981); Rousseeuw (1984, 1985); Rousseeuw & Molenberghs (1993); Falk (1997); Rousseeuw & van Driessen (1999); Friebel et al. (2010); Hubert et al. (2012). I tested several of them, which are mainly based on two different approaches.

The first group consists of algorithms that use the median or the median absolute deviation to estimate location and covariance matrix. Median and median absolute deviation are important quantities in statistics, because they have the highest possible breakdown point of 0.5. The breakdown point of an estimator is a basic measure of its robustness and defines the smallest amount of contamination that may cause an estimator to be incorrect (Huber, 1981; Donoho & Huber, 1983). This high breakdown point is in stark contrast to the arithmetic mean in equation (2.5) and the covariance matrix in equation (2.6), as they have a breakdown point of 0. The median absolute deviation (MAD) is used as a robust scale estimator for a univariate sample $\mathbf{Y}$, with $\mathbf{Y}$ being e.g. a column of our data matrix:

$$MAD = median\left|\mathbf{Y} - median(\mathbf{Y})\right|. \tag{2.7}$$

In statistics, the scale of a data set is defined as a measure of its dispersion. The conventional measures of scale are e.g. the sample variance or the sample standard deviation, which are both not robust and therefore highly influenced by outliers. However, a robust estimation of the variance σ can be derived from the median absolute deviation in equation (2.7) for independent and identically distributed data (Huber, 1981; Friebel et al., 2010):

$$\sigma = k * MAD \tag{2.8}$$

with k being a distribution dependent scale factor. However, MT data do not fulfil the assumption of independent and identically distributed data as they normally contain outliers. Therefore, equation

(2.8) can only be used as an initial guess of the variance. As shown later, this simple approach results in negative covariances indicating that it is not adequate for MT data.

The variance describes the dispersion of one variable. In contrast, the covariance matrix $\mathbf{C}_x$ consists of the variances of all p variables and all possible covariances Cov of a multivariate data set. Several equations for a robust covariance estimation based on the median absolute deviation are presented by e.g. Huber (1981) and Friebel et al. (2010). Most of these equations are related to a simple approach that was first proposed by Gnanadesikan & Kettenring (1972):

$$\mathbf{Cov}(\mathbf{Y}_i, \mathbf{Y}_j) = \frac{1}{4}\left(variance(\mathbf{Y}_i + \mathbf{Y}_j) - variance(\mathbf{Y}_i - \mathbf{Y}_j)\right) \tag{2.9}$$

with $\mathbf{Y}_i, \mathbf{Y}_j$ being two different columns of the data matrix $\mathbf{X}$.

An alternative covariance formulation entirely based on the median absolute deviation uses the comedian matrix Com (Falk, 1997):

$$\mathbf{Com}(\mathbf{Y}_i, \mathbf{Y}_j) = median((\mathbf{Y}_i - median(\mathbf{Y}_i))(\mathbf{Y}_j - median(\mathbf{Y}_j))). \tag{2.10}$$

All mentioned covariance equations based on the median absolute deviation are straightforward to implement and were tested for various MT data sets containing different noise contaminations. In general, all of these approaches were able to deal with MT data sets consisting of several thousands of events in an acceptably fast computation time. Furthermore, the location and the covariance matrix could be estimated well by these algorithms. However, all of these coordinate-dependent estimators have one large drawback, already pointed out by Rousseeuw & Molenberghs (1993) and Falk (1997): The estimated covariance matrix is not necessarily positive (semi-)definite.

By definition, a positive (semi-)definite matrix $\mathbf{A}$ obeys the formulation $\mathbf{x}^T\mathbf{A}\mathbf{x} \geq 0$ with $\mathbf{x}$ being any arbitrary non–zero vector. If the covariance matrix $\mathbf{C}_x$ in equation (2.1) is a positive (semi-)definite matrix, this formulation ensures that the squared MD is always positive. Although the empirical covariance matrix in equation (2.6) is always positive (semi-)definite, this does not have to hold for covariance matrices estimated by the median absolute deviation. Therefore, it is possible to obtain negative squared MDs, which makes it impossible to use them for data confinement.

Non-positive definite covariance matrices could be observed for all tested median absolute deviation algorithms, especially (i) for long period data that are typically computed on the basis of very few events or (ii) for severely disturbed periods, e.g. around the fundamental frequency of power grids. The occurrence of non-positive definite covariances is a hint that the simple robust statistics, relying

on median and median absolute deviation, are not applicable to MT data as the assumption of independent and identically distributed data might be violated.

Rousseeuw & Molenberghs (1993) and Maronna & Zamar (2002) introduced several approaches to transform non-positive (semi-)definite matrices into positive ones, which could be used for the covariance matrices formulated above. However, in the framework of this thesis this approach was not pursued as the alternative option provided similarly good results and its computational cost was only insignificantly higher. Furthermore, the second approach was explicitly designed for data sets that contain outliers.

The second group of robust location and covariance matrix estimators consists of highly robust estimators, e.g. the minimum volume ellipsoid (MVE) and the minimum covariance determinant (MCD) algorithm from Rousseeuw (1984, 1985). In the framework of this thesis, only the MCD algorithm was studied in detail because of the availability of fast algorithms and its ability to deal with larger data sets. Since Rousseeuw & van Driessen (1999) developed a fast MCD algorithm, this algorithm is most commonly used to calculate robust MDs.

The fundamental concept of the MCD algorithm is to calculate the location and the covariance matrix only from a subset H of all data consisting of h events. This subset should be ideally noise-free. It is chosen in such a way that it minimises the determinant of the covariance matrix as the determinant is a measure of the volume of a distribution (Basu & Ho, 2006). The larger the determinant, the more dispersed the data points. The aim of the algorithm is to find h points that are focused around the data centre. The general MCD algorithm consists of three key steps that will be explained in more detail below: (i) selection of various initial subsets, (ii) a kernel routine called concentration step (C-step) and (iii) selection of the final subset having the lowest covariance determinant and calculation of location and covariance from this subset. The size h of the subset is usually determined by the user and remains fixed during the procedure. The highest possible breakdown point of the MCD algorithm is achieved, when $h = (n + p + 1)/2$.

The centrepiece of this algorithm is the iteratively applied C-step. C-steps are applied individually for a large amount of initial subsets H_i to obtain a more accurate approximation to the MCD. Rousseeuw & van Driessen (1999) proved that C-steps converge in a finite number of steps. They showed that the determinant of the covariance matrix of the new subset is always lower or equal to the covariance determinant of the previous step, with equality only if the covariance matrices of both steps are identical.

Each C-step can be divided into three parts: (i) from a large amount of initial $h-$subsets H_i estimates

$\bar{\boldsymbol{x}}_{old}$ for the location and $\mathbf{C}_{old}$ for the covariance can be computed. The distances $d_{old}(i),\ i = 1, \ldots, n$ can be calculated by using $\bar{\boldsymbol{x}}_{old}$ and $\mathbf{C}_{old}$ in equation (2.1), (ii) these distances are sorted in an ascending order and the h events with the smallest distances are chosen to form a new subset and (iii) new estimates of the location and covariance matrix are calculated from the subset. C-steps are repeated until convergence is reached.

There is no guarantee that the result of the C-steps represents the global minimum; thus the MCD algorithm from Rousseeuw & van Driessen (1999) starts from many initial, randomly determined subsets. Furthermore, the raw MCD solution has a low statistical efficiency. In statistics, an estimator is regarded as efficient if it possesses the smallest variance among all possible estimators meaning that it achieves the lower bound on the Cramér-Rao inequality (Kendall & Buckland, 1957). The statistical efficiency of the MCD algorithm is increased by applying a one-step weighting approach, which uses only data points with a small distance value.

2.3.2 Implementation of a deterministic MCD algorithm

The MCD algorithm for a robust MD calculation used in the framework of this thesis is a further development of the deterministic MCD algorithm (Hubert et al., 2012). Hubert et al.'s (2012) algorithm was originally implemented in Matlab and is part of LIBRA, the Matlab Library for Robust Analysis (Verboven & Hubert, 2010). I programmed key parts of this routine into the software package EMERALD using C++. The work flow of the MD criterion with the MCD algorithm is visualised in Figure A.2.

The MCD algorithm utilises the data matrix $\mathbf{X}$ (section 2.1), which consists of n rows representing the observations, i.e. number of events for the examined period, and p columns representing the variables, i.e. real and imaginary parts of transfer functions. The deterministic MCD algorithm (Hubert et al., 2012) starts with a standardisation (normalisation) of this data matrix. Thus the algorithm becomes scale and location equivariant and results in the standardised matrix $\mathbf{Z}$. Each column of the data matrix is separately standardised by subtracting its coordinate-wise median $\boldsymbol{\mu}$ and dividing it by a robust scale estimator $\boldsymbol{\sigma}$:

$$\mathbf{Z} = \frac{\mathbf{X} - \boldsymbol{\mu}}{\boldsymbol{\sigma}}. \tag{2.11}$$

In the following, the three main steps of the implemented MCD algorithm will be explained, differences to the general approach from Rousseeuw & van Driessen (1999) are shown and an additional estimator is introduced.

- Each MCD algorithm starts by selecting various initial subsets. This is important as there is no guarantee that the final results of the iteratively applied C-steps represent the global minimum of the MCD objective function. The original MCD algorithm (Rousseeuw & van Driessen, 1999) takes many initial subsets and starts by drawing random subsets of size $p + 1$. However, as the computation time of the algorithm is proportional to the number of initial subsets, we deal with increased time with increased number of initial subsets. Furthermore, Rousseeuw & van Driessen's algorithm (1999) is not permutation invariant. It relies on randomly chosen subsets and therefore the final result depends on the order of the observations in the data set. This behaviour is actually not desired for a processing algorithm, as it should always come to the same results to ensure repeatability. Therefore, the modified deterministic MCD algorithm (Hubert et al., 2012) was implemented into EMERALD. Thereby, the algorithm starts only from six well-chosen statistical subsets determined by different estimators. The algorithm becomes permutation invariant and faster due to the application of always the same estimators instead of randomly drawn subsets. For each of the six estimators an initial estimate of the covariance or correlation matrix of the standardised data are computed. In my work, I use the same six estimators without any additional modification. Interested readers will find detail how to compute these statistical based estimators in the paper by Hubert et al. (2012).

  Tests of initial location and covariance estimates from these six estimators revealed that they work quite well for MT data. However, especially the initial location estimates are often very similar and close to the final estimated value. Because of this observation, I saw the need of a more independent seventh estimator. The new estimator is significantly different from the other six, mathematical estimators to ensure that its location differs from the other location estimates. Furthermore, it makes use of some physical relationships inherent to MT data. My estimator utilises the final location and covariance estimates from a previously processed, adjacent period as an initial guess to form a subset. The reasoning is based on the physics of induction processes: MT transfer functions vary only smoothly with period, i.e. slightly increased induction volume. Tests showed that this novel estimator could increase the variability of start values as it often results in totally different location and covariance estimates. If the results of the adjacent periods were biased or incorrectly estimated, this does not automatically lead to a wrong result for the currently examined period as the majority of the current data points have to confirm the solution to keep it. Therefore, it is relatively simple to get rid of false initial estimates during the iterative process and the novel implemented estimator cannot have a negative impact.

- In the second step, the raw MCD solution for each subset is separately computed by applying C-steps until convergence is reached.

- Finally, the subset having the lowest covariance determinant is selected and location and covariance are computed from this subset. This raw MCD solution can have a breakdown value up to 50 %; however, its efficiency for Gaussian distributions is low. Therefore, a one-step weighting is applied to increase the statistical efficiency and retaining the high breakdown point at the same time. Processing results obtained by the raw and re-weighted solutions were compared for many stations with different noise contaminations. In general, the processing results obtained by the re-weighted MCD solution are slightly superior, i.e. resulting in smoother transfer functions. However, in most cases the difference was negligible.

The final MD criterion consists of three main steps: (i) the data matrix $\mathbf{X}$ is formed by computing all single event transfer function components, (ii) the location and covariance matrix are robust estimated by using the deterministic MCD algorithm and (iii) a MD value is computed for each single event. If the single event MD is lower or equal to a given threshold, then the single event is used for the further processing. Otherwise, it is rejected and not used in the subsequent stacking process.

The application of the MD criterion and the associated MCD algorithm requires some key parameters. These key parameters are (i) the size h of the subset H, which is chosen to estimate the location and the covariance matrix with $\frac{n+p+1}{2} \leq h \leq n$, (ii) the maximum number of allowed C-steps and (iii) the critical distance (threshold) above which events are characterised as outliers or noise affected data. The first two parameters are needed within the MCD algorithm. Both of them were hard-wired in the algorithm after several tests. Consequently, users do not have to take care about a correct choice of these two parameters. Only the third parameter has to be selected by the user as this parameter will depend on the EM noise characteristics of the MT data set.

The first of the three key parameters determines the size h of the subset H used for the MCD algorithm. This size can vary between the smallest possible subset with $h = \frac{n+p+1}{2}$, which represents the case where the algorithm has its highest breakdown point, and the theoretically maximal subset with $h = n$. This last case is not used, as it results in the normal arithmetic mean and the empirical covariance matrix. Instead of this value, the size is often set to $h = 0.75n$, assuming that the data are contaminated by less than 25 % noise. An $h-$ value of $0.75n$ is a good compromise between retaining a high breakdown point and having a good statistical efficiency (Rousseeuw & van Driessen, 1999; Hubert et al., 2008; Hubert & Debruyne, 2010; Verboven & Hubert, 2010; Hubert et al., 2012). However, a smaller h should be used in presence of a significant (e.g. > 25 %) or unknown amount of noise. Three different $h-$values were tested for a variety of MT stations with $h = \left(\frac{n+p+1}{2}, 0.6n, 0.75n\right)$. Only small differences were observed for stations that already have an acceptable data quality. Larger effects were noticed for stations that are contaminated to a larger extent by noise. Unfortunately,

nowadays many data sets are highly affected by man-made noise and therefore contain a large amount of distorted events at least for some periods. For these data sets, the lower limit for h yields the best results due to the higher robustness of the location and covariance matrix estimates against outliers and noise. The effect of different $h-$values is exemplarily shown in Figure 2.2 and 2.3 for a station from South Africa.

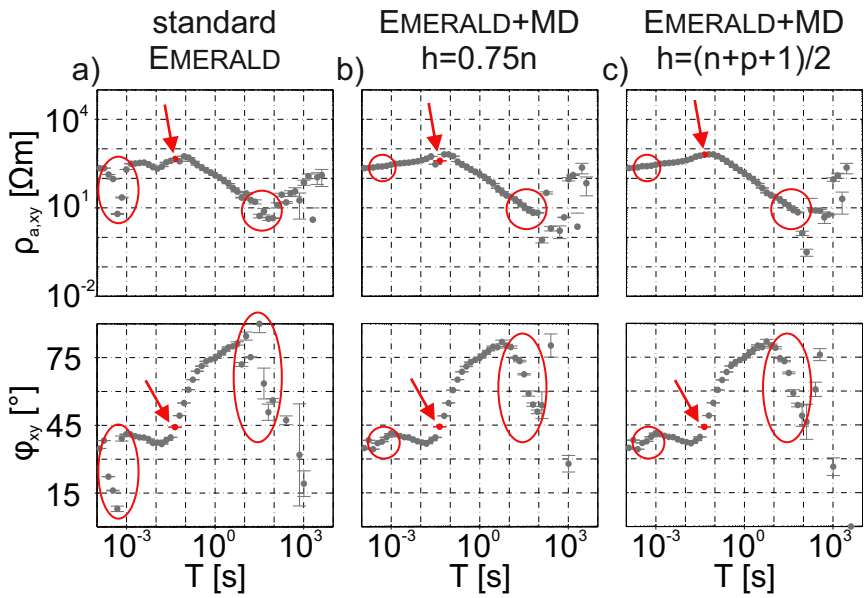

Figure 2.2: Apparent resistivity and phase curves of station SA-415 for different subset sizes of the MCD algorithm. a) Results using standard EMERALD processing. b) Results using EMERALD+MD processing with $h = 0.75n$. c) Results using EMERALD+MD processing with $h = (n + p + 1)/2$. Processing results for periods highlighted by the red ellipses and circles could be improved using EMERALD+MD processing. The period of $T = 1/22.63\,s$ is highlighted and further investigated in Figure 2.3.

The standard EMERALD processing result is compared with two EMERALD+MD processings using $h-$values of $0.75n$ and $\frac{n+p+1}{2}$ (Fig. 2.2). Both EMERALD+MD processing results (Figs. 2.2b & c) show an improvement in comparison with the standard EMERALD processing (Fig. 2.2a) for many periods indicated by the red ellipses and circles. Focusing on the highlighted (red) Z_{xy} data points in apparent resistivity (upper row) and phase (bottom row) differences are visible; while the choice of $h = 0.75n$ improves the results slightly compared to the standard EMERALD processing results; $h = \frac{n+p+1}{2}$ retrieves the best result. The period of $T = 1/22.63\,s$ is chosen to emphasise the difference by showing the histograms of all single events for the three different processing settings and for the selected period in Figure 2.3.

The selected period contains coherent noise, which results in two data clusters. To conclude which cluster belongs to natural MT excitation, robust estimates for the adjacent period of $T = 1/16\,s$ are used as a reference, represented by the black triangle in Figure 2.3b and c.

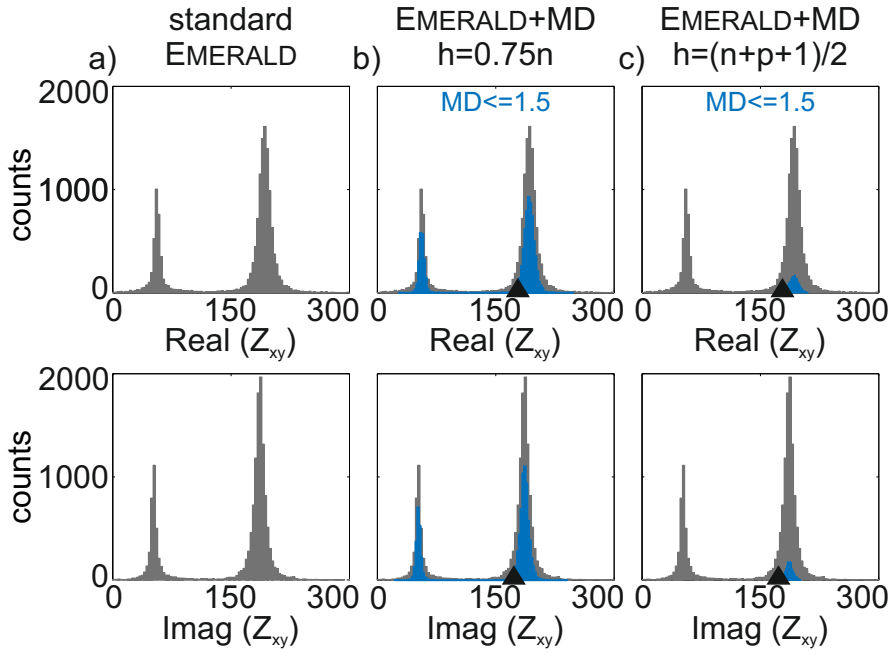

Figure 2.3: Histograms of real and imaginary part of Z_{xy} for all events of $T = 1/22.63\,s$. Extreme outliers are not displayed as they can be removed from the standard robust statistics as well as from the MD criterion. a) The histograms of the selected period show, that the data are separated into two clusters (caused by MT signal and EM noise). b) The histograms of all events in grey colours are overlayed by the blue-coloured single events, which are accepted by the MD criterion with a threshold of 1.5. c) Similar to b) but with a different h - value. The black triangles symbolise real and imaginary parts of the robust estimated Z_{xy} component for the adjacent period of $T = 1/16\,s$. In b) data from the desired MT signal (right cluster) as well as from EM noise (left cluster) are used for a robust MD calculation leading to disturbed results in contrast to c) where only data of the desired MT signal are used.

This period is not disturbed by noise and was taken as the seventh estimator within the MCD algorithm. From this follows, that the cluster to the right belongs to the desired MT signal. Distributions of all available events are overlayed by the blue-coloured distributions of the data used for the robust MD calculation (Figs. 2.3b & c). In Figure 2.3b, 75 % of all available data are used for the calculation of robust location and covariance matrix, assuming that the data are contaminated by less than 25 % with noise. Visual inspection reveals, that this assumption is violated and a kind of masking effect is observed, meaning that there is no subset H, which is not disturbed and influenced by noise. Therefore, the final subset H contains a moderate amount of noise, which influence the location and covariance matrix estimates. Consequently, the derived MDs are no longer able to detect all noise affected data and events of both clusters - originating from noise and from the desired MT signal - are used for the stacking process. In contrast, if location and covariance matrix are estimated from only half of the data (Fig. 2.3c), the MD criterion is able to identify the correct distribution. Therefore, the derived MD values are useful to confine the data to a noise-free subset. The processing result of the standard EMERALD processing for this period is similar to the result in 2.2c indicating that the

standard EMERALD processing has a breakdown value greater than 25 %.

These kind of comparisons with different data sets motivated me to always assume the worst case and severest noise contamination and therefore I fixed the value of h to $\frac{n+p+1}{2}$.

The second parameter is the maximum number of C-steps applied to each initial estimator. Ideally, C-steps should be applied until convergence is reached. However, Rousseeuw & van Driessen (1999) introduced a limited number of C-steps in their fast MCD algorithm in the interest of speed, especially for large data sets. Hubert et al. (2012) followed this example in their deterministic approach and fixed the maximum number of C-steps to 100. The same limit is used for the MCD algorithm implemented in EMERALD.

To evaluate this maximum, tests with different stations and data qualities were conducted. These tests sugggest that the implemented MCD algorithm needs much less C-steps even with measured MT data that are highly contaminated with noise. The maximum value of 100 C-steps was never reached for all tested stations, so that this limit seems to be sufficient. A moderate increase of applied C-steps is observed for short periods in comparison to long periods due to the larger number of available events.

The third key parameter for the MD criterion is the critical distance to characterise outliers and noise affected data. For this purpose, a threshold has to be chosen. The optimal threshold depends on the data quality and therefore has to be determined for each data set and station by the user. As a starting point the values in table 2.1 (page 19) can be used.

In the case of EMERALD processing, where the transfer functions are estimated by solving a bivariate equation, a useful interval for thresholds could be $[1, 4]$. The upper boundary is derived from $\alpha = 0.995$ and the lower boundary was found by several tests. As an alternative, the threshold can be calculated using Chebyshev's inequality in Section 2.2. Due to the independence of an assumed data distribution, this always will result in higher values than the computed thresholds by using the quantiles of the χ^2-distribution in table 2.1.

The influence of different thresholds is exemplarily displayed for apparent resistivity and phase curves of the Z_{xy} component (Fig. 2.4) and as histograms (Fig. 2.5) for station V-027 from Venezuela. All three EMERALD+MD processings (Figs. 2.4b - d) are superior to the standard EMERALD processing (Fig. 2.4a). However, the most effective threshold in this example is 1.8, which results in an almost completely smooth apparent resistivity and phase curve except for only a few periods.

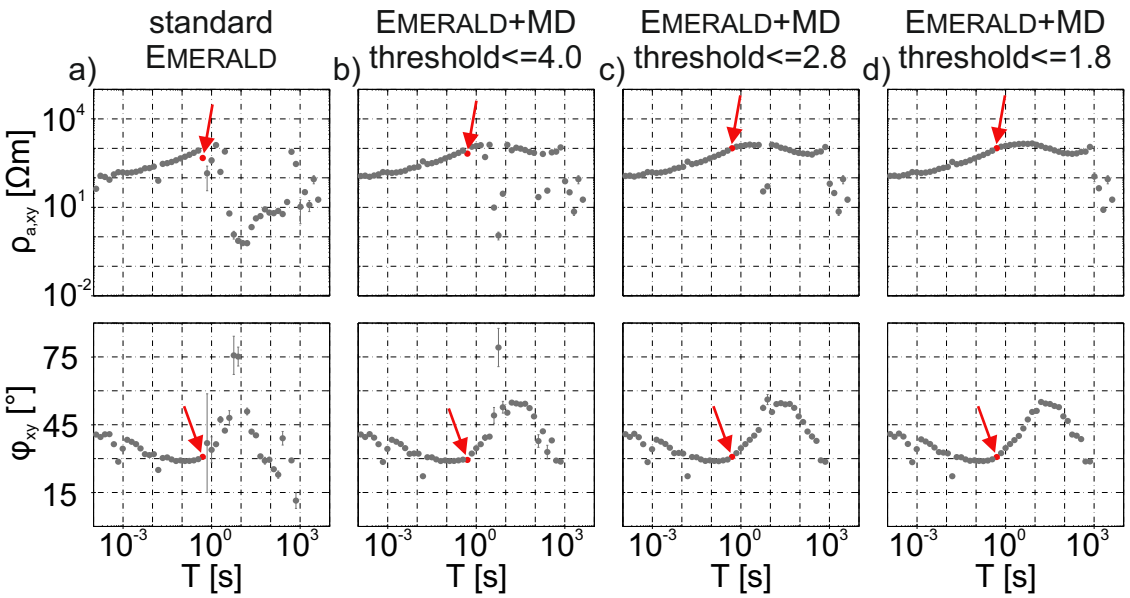

Figure 2.4: Apparent resistivity and phase curves of station V-027 for different thresholds for the MD criterion. a) Results of standard EMERALD processing. Results of EMERALD+MD processing with a threshold of b) 4.0, c) 2.8 and d) 1.8. The period of $T = 0.5\,s$ is highlighted and further investigated in Figure 2.5.

The period of $T = 0.5\,s$ is selected to further investigate the influence of the different thresholds (Fig. 2.5). The colour of each single event represents its smoothed likelihood (Eilers & Goeman, 2004).

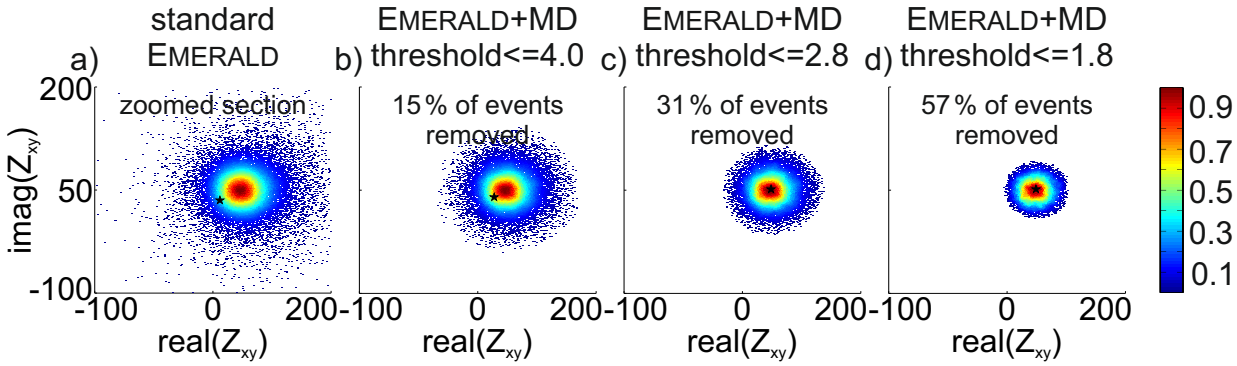

Figure 2.5: Scatterplots for a period of 0.5 *s*. Each single event is colour-coded with its smoothed likelihood (red ≙high, blue ≙ low likelihood). The black asterisks represent the final processing results for the selected period. a) Data distribution without application of the MD criterion. In b) to d) the MD criterion is applied with decreasing thresholds leading to a smaller and more focused data distribution.

In standard EMERALD processing (Fig. 2.5a) accepted data are distributed in one large cluster; a minor fraction of data scatters around this cluster and hampers the transfer function estimation (final result: black asterisk in Fig. 2.5a). Obviously, the calculated transfer function does not correspond well with the area of maximal likelihood of occurrence. In Figure 2.5b to d the threshold is successively

decreased and consequently more data are rejected, which results in a more focused cluster. For the selected period a threshold of 2.8 seems to be sufficient to yield a good processing result. This corresponds with the processing results shown in Figure 2.4. However, for some periods, e.g. 5.6 s or 8 s a smaller threshold is needed to obtain the best possible result. This example illustrates that the optimal threshold can vary even for different period data of one station. Therefore, the user should carefully examine the optimal threshold for each data set and station. If the threshold is too small, the transfer function estimation becomes unstable due to the too small amount of accepted events. Contrary, if the threshold is chosen too large, the MD criterion will not be able to remove a sufficient amount of noise affected data. This can also result in incorrect transfer functions as the remaining noise might still hamper the transfer function estimation. It is the task of the user to find a good compromise between these two boundary conditions by evaluating the smoothness of the different transfer functions by visual inspection.

2.4 Combination of the Mahalanobis distance criterion with other data selection criteria

The MD criterion is a purely statistical criterion and hence it does not distinguish between physically reasonable and non-physical or near-field EM noise data. Therefore, the MD criterion will always favour clusters comprising of the majority of data points. The processing routine for the robust transfer function estimation within EMERALD includes already physically motivated pre-stack data selection criteria. These criteria are able to remove data points that disagree with an assumed physical model. For this reason, it is advisable to apply these data selection criteria prior to the application of the MD criterion.

The most commonly used data selection criterion in the EMERALD processing is the coherence criterion based on the bivariate quadratic coherence in equation (1.14). The coherence criterion removes events with a small coherence. The influence of the coherence criterion on the transfer function estimation is exemplarily demonstrated in Figure 2.6 and 2.7. Apparent resistivity and phase curves (Fig. 2.6) show some scatter at short periods, in the dead band and for long periods. Best processing results are obtained by a combination of coherence and MD criterion prior to the standard EMERALD processing. Especially the period range of $T > 1/512\,s$, that includes the highlighted data point, and the dead band can be improved for this station.

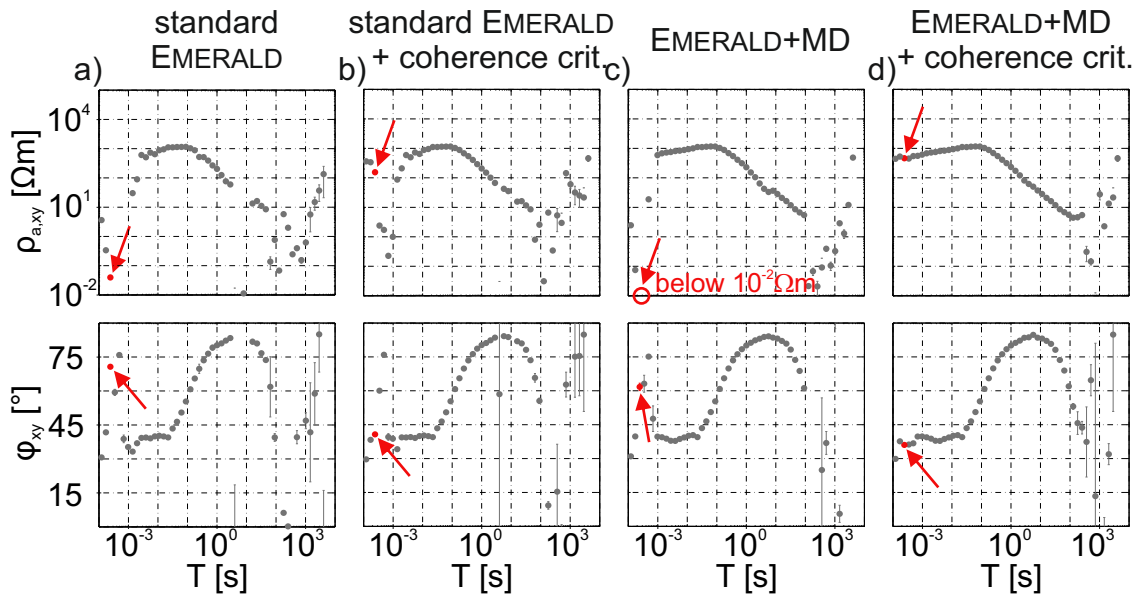

Figure 2.6: Apparent resistivity and phase curves of station SA-217 showing the influence of the coherence criterion. a) Results of standard EMERALD processing without applying the coherence criterion. b) Results of standard EMERALD processing with a coherence threshold of 0.9. c) Results of EMERALD+MD processing without using the coherence criterion. d) Results of EMERALD+MD processing applying an additional coherence threshold of 0.9 prior to the application of the MD criterion. The period of $T = 1/4096\,s$ is highlighted and further investigated in Figure 2.7.

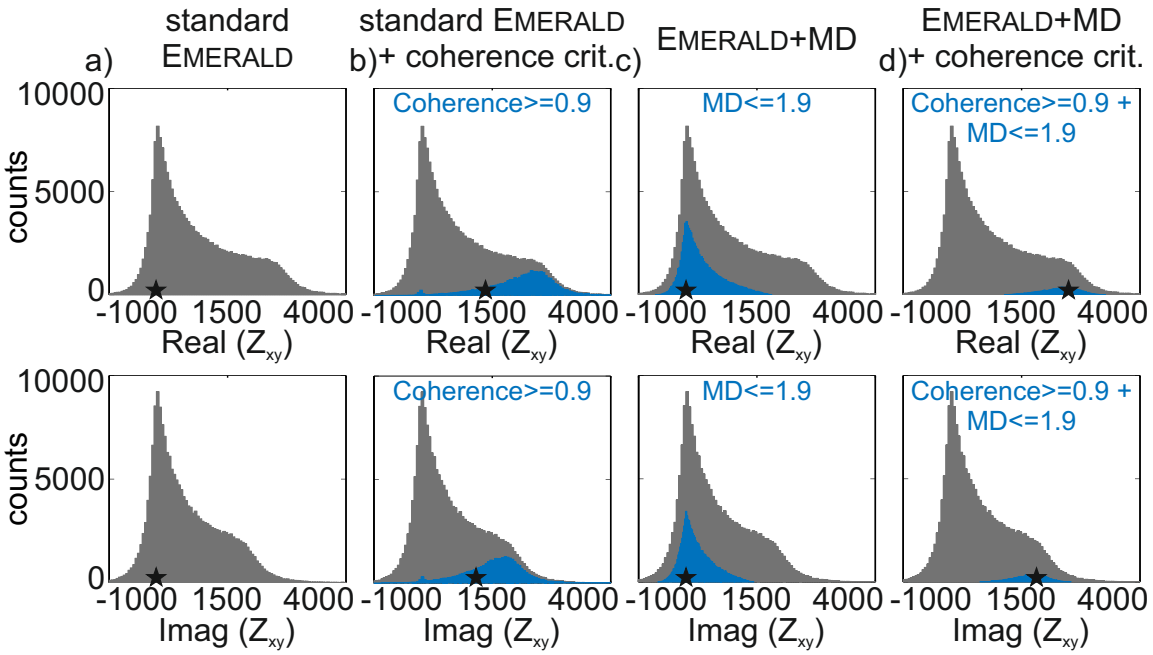

Figure 2.7: Histograms of real and imaginary part of Z_{xy} for station station SA-217 and $T = 1/4096\,s$. a) The grey-coloured distributions represent all single events, which are used for the standard EMERALD processing. b) The grey-coloured distributions of a) are overlayed by the blue-coloured distributions, which represent all events with a coherence greater or equal to 0.9. c) The blue-coloured distributions represent all events with a MD lower or equal to 1.9. d) The blue-coloured distributions symbolise all events with a coherence greater or equal to 0.9 and a MD lower or equal to 1.9. The black asterisks represent the final processing result. Best results are obtained by the combination of coherence and MD criterion.

The period of $T = 1/4096\,s$ (highlighted in Fig. 2.6) is selected to give an extreme example of the efficiency of the coherence criterion. The grey-coloured histograms (Fig. 2.7) display all available single events, indicating a broad data distribution. The transfer function estimated by the standard EMERALD processing (black asterisk in Fig. 2.7a) coincides with the maximum of the distribution. However, the distribution plots for standard EMERALD processing plus coherence criterion (Fig. 2.7b) reveal that most of the events with obviously low coherence leading to the poor final impedance estimate for standard EMERALD processing (Fig. 2.6a). Only the blue-coloured events have a coherence value greater or equal to 0.9 (Fig. 2.7b). Events used for the EMERALD+MD processing without (Fig. 2.7c) and with (Fig. 2.7d) an additional coherence threshold are again displayed in blue. Application of the MD criterion without coherence criterion results in a completely misleading transfer function. The MD criterion as a purely statistical approach will concentrate on the majority of data without recognising that these points are physically not well-behaved. The data points originating from natural MT signal are almost completely removed in this example. In contrast, the additional application of the coherence criterion prior to the MD criterion leads to the correct transfer function (Fig. 2.7d).

Therefore, in most cases it is recommended to use the MD criterion in combination with the coherence criterion. If the coherence and MD thresholds are well chosen, the processing results can be significantly improved.

As an additional selection criterion, Weckmann et al. (2005) took advantage of the fact that MT phases typically lie in one quadrant in case of a simple subsurface conductivity structure. This so-called phase criterion was used as a brute force option due to a lack of other alternative criteria at that time. Although the application of the phase criterion together with the coherence criterion sometimes results in smoother apparent resistivity and phase curves, the underlying assumptions are often violated, resulting in seriously misleading final estimates. Due to its brute force and often unphysical nature, the phase criterion often destroys the underlying Gaussian distribution of the data by truncating the data distribution or eliminating large parts of the data.

This problem is illustrated by the comparison of standard EMERALD processing (Fig. 2.8a), standard emerald plus phase criterion (Fig. 2.8b) and the EMERALD+MD processing (Fig. 2.8c). Apparent resistivity and phase values reveal that the application of the phase criterion results only in slightly smoother curves, whereas the application of the MD criterion leads to a significant improvement.

The period of $T = 1/2048\,s$ is chosen to display scatterplots for an illustration of the effect of the phase and the MD criterion in Figure 2.9.

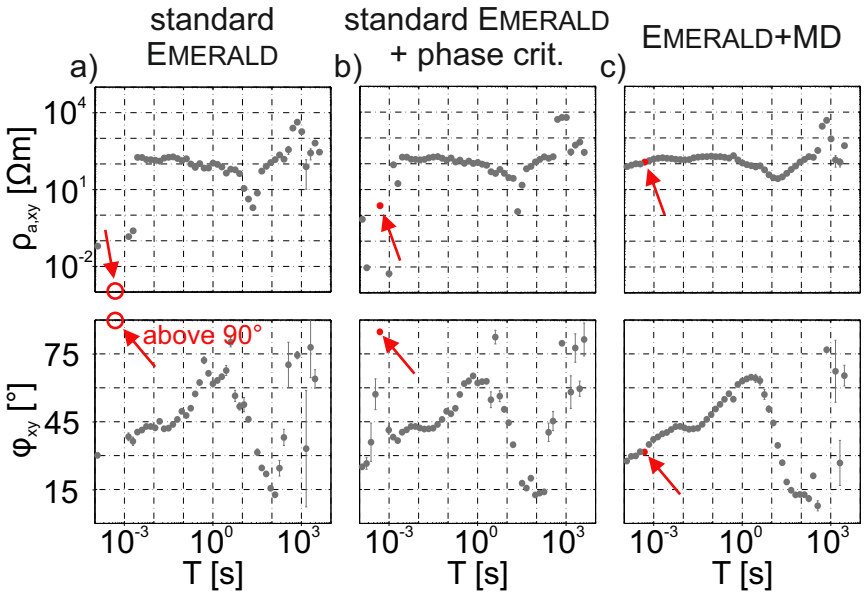

Figure 2.8: Apparent resistivity and phase curves of station SA-704 for different processing settings. All processings used a coherence threshold of 0.9. a) Results of standard EMERALD processing. b) Results of standard EMERALD processing using phase criterion. c) Results of EMERALD+MD processing. The period of $T = 1/2048\ s$ is highlighted and further investigated in Figure 2.9.

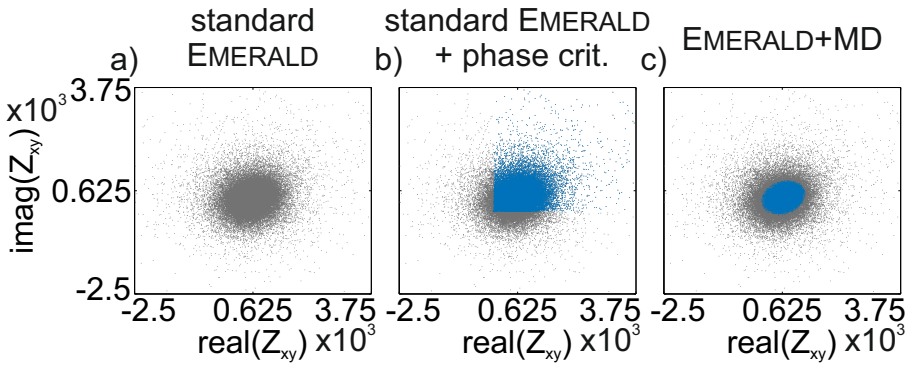

Figure 2.9: Scatterplots for a period of $T = 1/2048\ s$. a) Each single event is represented by a grey dot in the complex plane. b) Only the blue-coloured dots are accepted by the phase criterion representing a truncation of the original distribution. c) Only the blue-coloured dots are accepted by the MD criterion focusing on the centre part of the original distribution.

The effect of the different criteria can be seen in the colour-coded histograms of events (Fig. 2.9); in grey we have all events with coherences > 0.9, blue indicates the impact of the phase (Fig. 2.9b) and the MD criterion (Fig. 2.9c). The application of the phase criterion truncates the data distribution, which results in the poor and erroneous processing results. In contrast, the MD criterion only accepts events in the centre of the data distribution, preserves the Gaussian core and yields the desired estimate.

In summary, it is expedient to use a coherence threshold in addition to the MD criterion. However,

the usage of the phase criterion often results in erroneous estimates as the underlying distribution is truncated. Here, I recommend to use the MD plus coherence criterion instead.

2.5 Application of the Mahalanobis distance criterion to different data sets

In this section, I present processing results with different noise contaminations to show advances and limitations of the new criterion as well as to demonstrate under which circumstances improved results can be obtained.

2.5.1 Scattered distributions

Although station SA-509 has an overall fair data quality, standard EMERALD processing results are only acceptable in some period ranges. Applying the MD criterion in addition significantly improves the processing results and yields much smoother apparent resistivity and phase curves (Fig. 2.10).

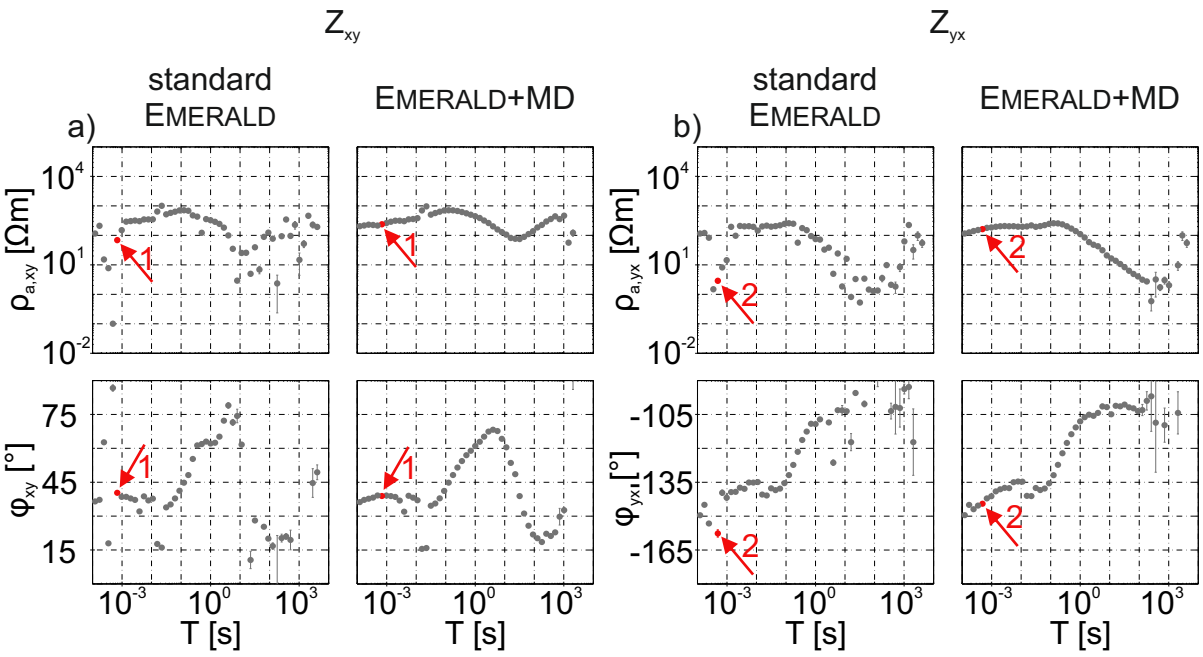

Figure 2.10: Apparent resistivity and phase curves of station SA-509 comparing standard EMERALD and EMERALD+MD processing for a) Z_{xy} and b) Z_{yx}. The periods of $T_1 = 1/1448\,s$ (Z_{xy}) and $T_2 = 1/2048\,s$ (Z_{yx}) are highlighted and further investigated in Figure 2.11.

Useful hints on how the EMERALD robust statistic in contrast to the addition of the MD criterion works can be obtained if we look at the distribution of events in the complex plain (Fig. 2.11).

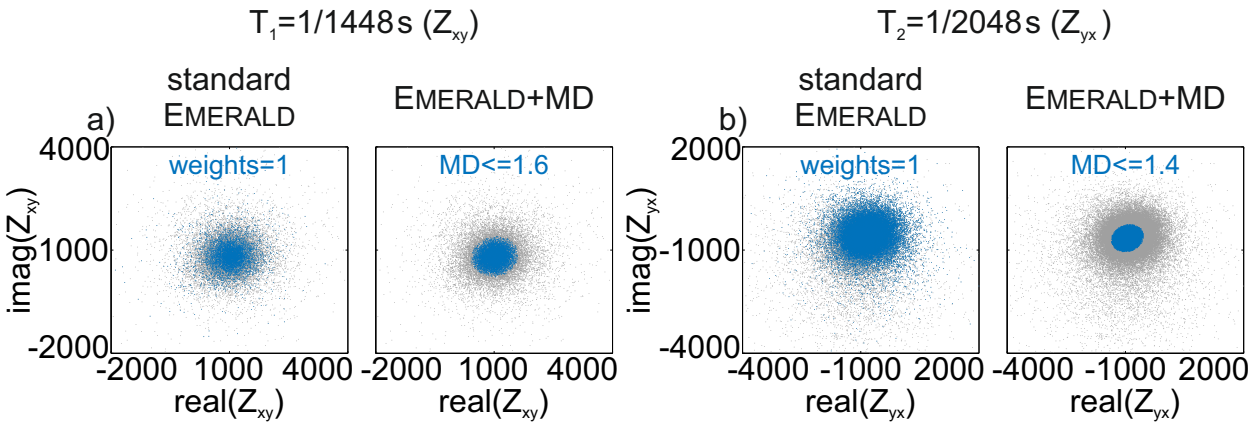

Figure 2.11: Scatterplots of station SA-509 for a) $T_1 = 1/1448\,s$ (Z_{xy}) and b) $T_2 = 1/2048\,s$ (Z_{yx}) comparing accepted events for standard EMERALD and EMERALD+MD processing. The distribution of all events is represented as grey dots in the complex plane and is overlain by all events with the highest possible weight in the robust stacking process (standard EMERALD) and all events below a certain MD value (EMERALD+MD), respectively. Application of the MD criterion leads to a more focused subset of data and rejection of scattered data points.

For the two highlighted periods (Fig. 2.10), scatterplots of both processings indicate that the events are arranged in one confined cluster. However, a minor fraction of data scatters around this cluster and hampers the transfer function estimation in the standard EMERALD processing. From the left-hand side scatterplots in Figure 2.11a and 2.11b, it is visible that some of the scattering events have the highest possible weight of 1 in the standard EMERALD processing and therefore these events have a large influence on the processing result. Applying the MD criterion prior to the stacking procedure removes these events as they have a larger distance to the actual data centre. This results in more focused clusters and consequently in smoother transfer functions.

This example demonstrates that the MD criterion is capable to improve processing results of stations affected by noise that scatters around the true MT distribution by removing scattering events and focusing on events located close to the data centre.

2.5.2 Two spatially separated distributions

Again, MT data of station D-308 show some distinct period ranges around $T = 1/32\,s$ and the dead band, where the apparent resistivity and phase values of both off-diagonal impedance tensor elements scatter (Fig. 2.12). Insight into the MT data properties can be obtained through the scatterplots for

the exemplary period of $T = 1/32\,s$ indicated in red in Figure 2.12.

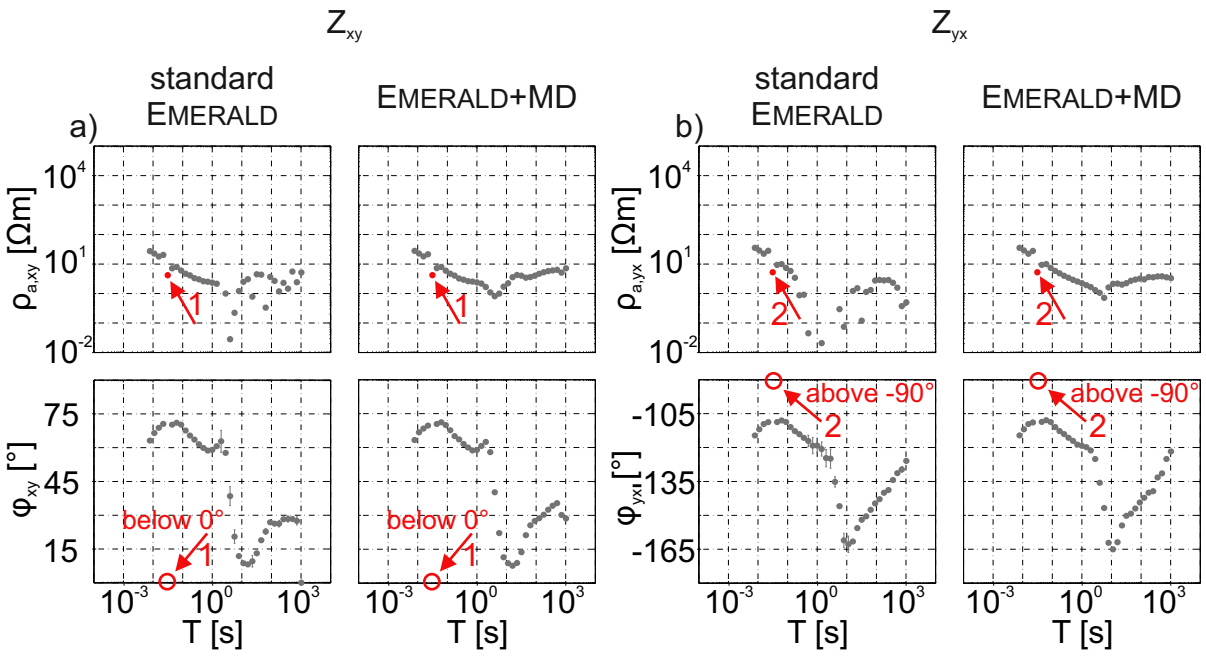

Figure 2.12: Apparent resistivity and phase curves of station D-308 comparing standard EMERALD and EMERALD+MD processing for a) Z_{xy} and b) Z_{yx}. The period of $T = 1/32\,s$ is highlighted for both components and further investigated in Figure 2.13.

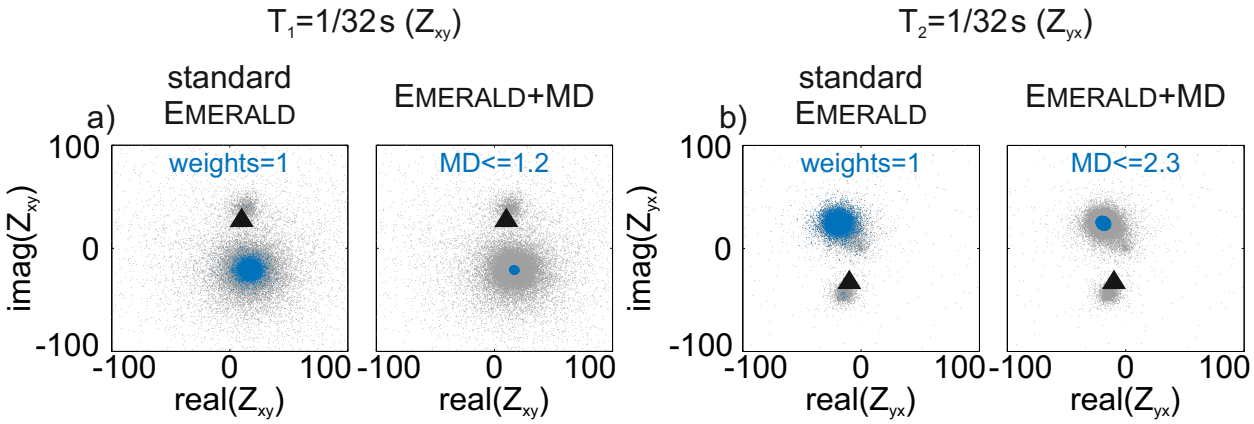

Figure 2.13: Scatterplots of station D-308 for a period of $T = 1/32\,s$ for a) Z_{xy} and b) Z_{yx} showing two spatially separated distributions. This example demonstrates that the standard EMERALD as well as the EMERALD+MD fail, if the majority of all events is caused by noise. The black triangles represent the processing results of the undisturbed adjacent period of $T = 1/22.63\,s$ indicating the distributions of the desired MT signal, which consist of the minority of all events (smaller cluster). The majority of all events (larger clusters) is caused by EM noise.

For both impedance tensor components, two spatially separated clusters exist (Fig. 2.13). The black triangles represent the processing results of the undisturbed adjacent period of $T = 1/22.63\,s$ and indicate the cluster belonging to the desired MT signal. For both components, the EM noise cluster

consists of the majority of data. As already mentioned, the MD criterion will fail in this case, as it will focus on the cluster representing the majority of events. Therefore, the application of the MD criterion will remove all desired MT signal and only events from noise sources will be accepted. To prevent the algorithm from choosing the wrong data cluster, interactive selection algorithms (e.g. from Weckmann et al., 2005) can be used to pre-select events by visual inspection of different physical parameters such as power spectra or polarisation directions of the magnetic or electric field. Another solution is to use the information of the undisturbed adjacent period as a-priori knowledge and to manually remove events corresponding to the EM noise cluster.

2.5.3 Two merged distributions I

With this example, I would like to illustrate what happened if we have two merged distributions of data. Two disturbed periods are selected (in Fig. 2.14a) as an example where a second cluster overlays the desired MT distribution to a large extent (Fig. 2.15).

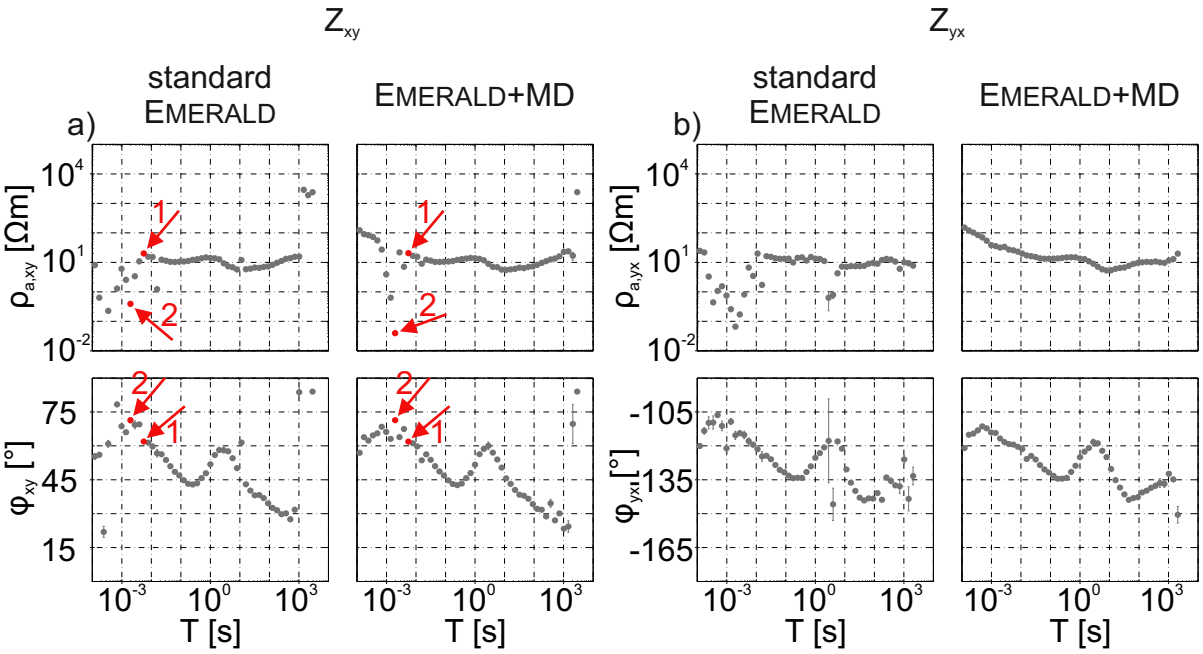

Figure 2.14: Apparent resistivity and phase curves of station V-117 comparing standard EMERALD and EMERALD+MD processing for a) Z_{xy} and b) Z_{yx}. The periods of $T_1 = 1/181\,s$ and $T_2 = 1/512\,s$ are highlighted for the Z_{xy} component and further investigated in Figure 2.15.

In Figure 2.15a, histograms of real and imaginary part as well as the corresponding colour-coded scatterplot are shown for $T = 1/181\,s$. This period seems to be affected by a small amount of EM noise as we observe a smaller second cluster (visible as long tail in the histograms). Events of this smaller

cluster can be removed by the MD criterion.

For $T = 1/512\,s$, the cluster of EM noise is larger than the cluster with the desired MT signal. Furthermore, both clusters are very close together and overlap to a large extent.

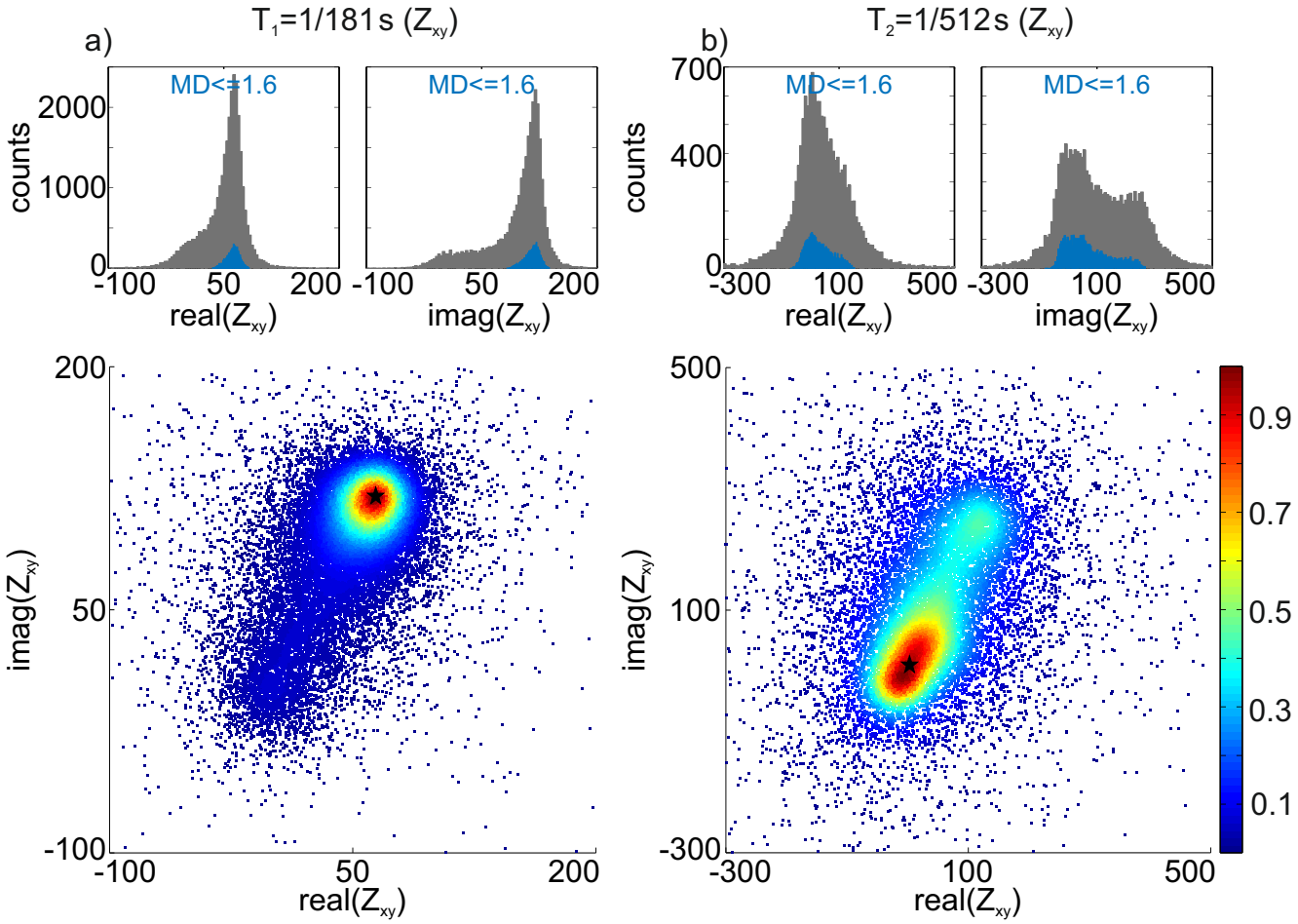

Figure 2.15: Histograms and colour-coded scatterplots of station V-117 for a) $T_1 = 1/181\,s$ and b) $T_2 = 1/512\,s$ for Z_{xy}. The colour-coded scatterplots complement the information of the histograms and visualise the distributions in the complex plane. The black asterisks represent the processing results for the selected periods. a) The distribution of all events (grey colour in histograms) is long tailed, best visible in the histogram of the imaginary part and in the corresponding scatterplot. The MD criterion completely removes events corresponding to this tail. b) Two merged distributions exist for this period, whereby the majority of data belongs to EM noise. Therefore, the MD criterion fails.

As a result, the application of the MD criterion mainly focus on events from the larger cluster, but interestingly also some points from the "true" cluster are considered due to the overlap of both clusters. However, the majority of all accepted events originates from noisy data and therefore we obtain disturbed MT transfer functions. Additional information is needed in such complicated situations, e.g. by visual pre-selection of events through different physical parameters such as the magnetic polarisation direction or by manual removing points based on information of an undisturbed adjacent

period.

2.5.4 Two merged distributions II

In general, station N-103 has a very high data quality; however, the MD criterion can only be improve the period range around the dead band (Fig. 2.16). The short periods $< 0.001\,s$ cannot be improved.

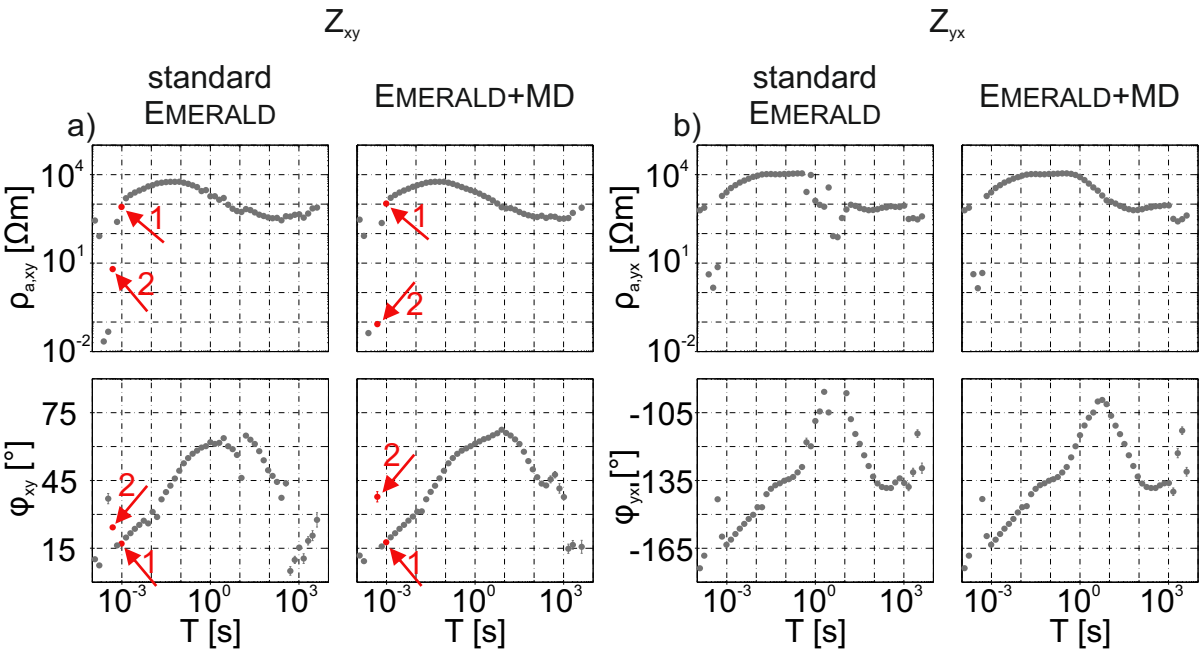

Figure 2.16: Apparent resistivity and phase curves of station N-103 comparing standard EMERALD and EMERALD+MD processing for a) Z_{xy} and b) Z_{yx}. The periods of $T_1 = 1/1024\,s$ and $T_2 = 1/2048\,s$ are highlighted for the Z_{xy} component and further investigated in Figure 2.17.

Histograms of real and imaginary part of Z_{xy} for the almost undisturbed period of $T = 1/1024\,s$ (Fig. 2.17a) seem to indicate one single cluster. However, a closer look reveals that some EM noise is included in the data in form of a bias, indicated by the red circle in the scatterplot. This noise is also visible in the histogram of the real part as a small peak on the left-hand side. Nevertheless, this noise only affects a small number of events and can be removed by the MD criterion.

In contrast, the period of $T = 1/2048\,s$ (Fig. 2.17b) is more severely affected by the same noise. At first glance, the data are again distributed in one single cluster. But the colour-coded scatterplot reveals that the desired MT distribution is overlain by a second distribution. Most of the events originate from this second distribution leading to the poor results for the standard EMERALD processing. In this case, application of the MD criterion impairs the situation even more. The MD criterion as a statistical criterion always focuses on the cluster consisting of the majority of points. If the majority

of points is caused by EM noise, it is possible that the MD criterion removes the desired MT signal completely or keeps events of both distributions. This latter case occurs when the clusters have about the same size or are merged to a significant extent as in the shown example. The MD criterion fails to remove noise, because the derived location and covariance matrix are estimated as an average over all clusters and therefore distort the MD values in a way that events originating from noise do not necessarily have a high MD value. In this case, a pre-selection of events either by visual inspection of different physical parameters or by using additional information of undisturbed adjacent periods is necessary to ensure that the majority of all events is well-behaved.

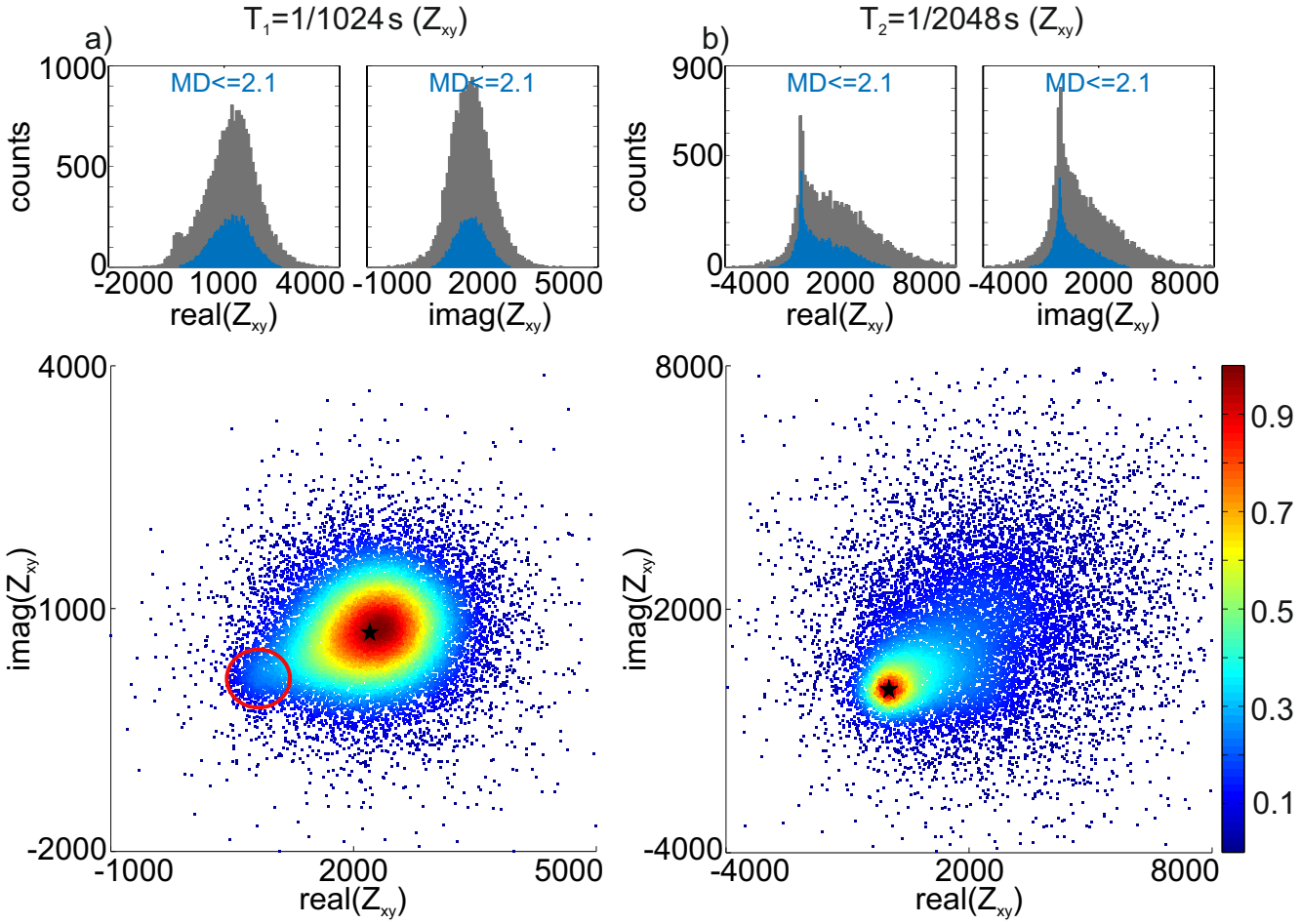

Figure 2.17: Histograms and colour-coded scatterplots of station N-103 for a) $T_1 = 1/1024\,s$ and b) $T_2 = 1/2048\,s$ for Z_{xy}. The black asterisks represent the processing results for the selected periods. a) A minor fraction of the data belongs to a second distribution (red circle) caused by EM noise and can be removed by the MD criterion. b) The majority of all events belongs to the EM noise distribution indicated in a) and overlays the desired MT distribution.

2.5.5 Application to remote reference processing

In addition to all previous examples, the MD criterion can also be applied to remote reference processing. However, in many tests the single site results using EMERALD+MD processing are similarly good or even better than using the standard EMERALD remote reference processing. This is a positive side-effect as often a reference station does not have sufficiently clean data or had some kind of technical problems so that we do not rely strongly on the remote reference processing to obtain acceptable data quality. Although, the existence of a true remote station is still essential in many cases to enhance processing results like in my next example (Fig. 2.18).

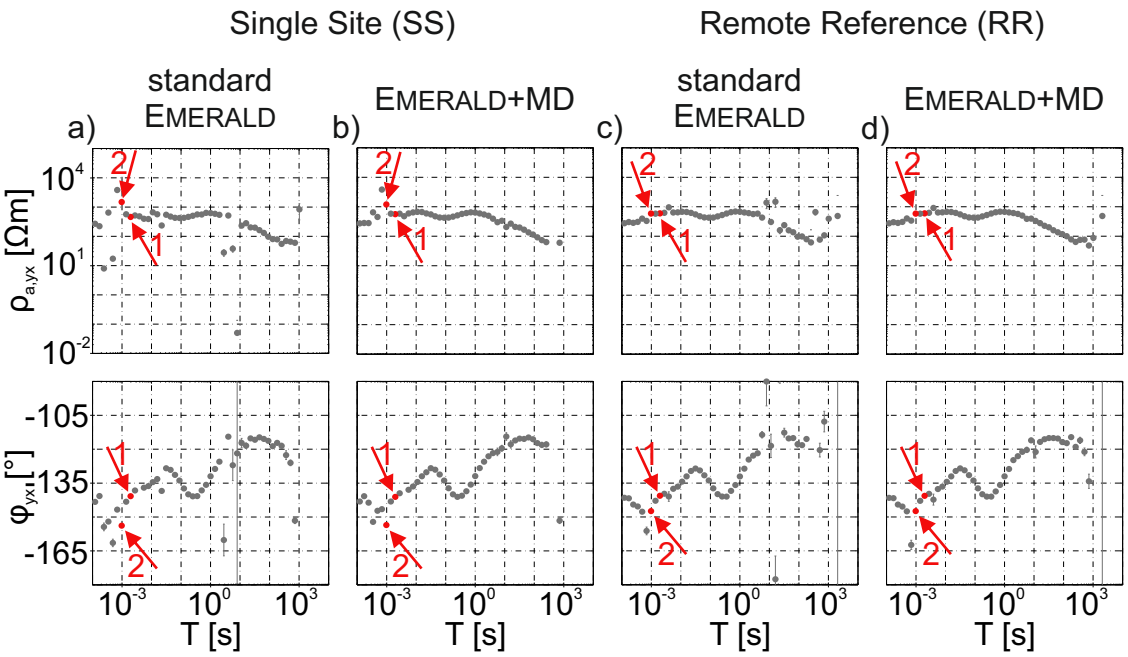

Figure 2.18: Apparent resistivity and phase curves of station T-420 (with T-502 as remote station) for a-b) single station (SS) and c-d) remote reference (RR) processing comparing standard EMERALD and EMERALD+MD processing for Z_{yx}. The periods of $T_1 = 1/512\,s$ and $T_2 = 1/1024\,s$ are highlighted and further investigated in Figure 2.19.

Here, we compare standard EMERALD and EMERALD+MD processing for single site (Figs. 2.18a & b) and remote reference (Figs. 2.18c & d) processing. The EMERALD+MD single site processing (Fig. 2.18b) improves especially the dead band between 1 s and 10 s in contrast to the standard EMERALD single site processing (Fig. 2.18a). However, it fails to improve the short periods $< 1/1000$ s. The standard EMERALD remote reference processing (Fig. 2.18c) can improve the short periods, but fails to improve the period range around 10 s. The best result can be obtained by applying the MD criterion for the remote reference processing (Fig. 2.18d).

The period $T = 1/512\,s$ was selected to illustrate under which circumstances both standard EMERALD

remote reference and EMERALD+MD single site processing can yield improvements. Normally, remote reference processing is applied to remove bias due to noise in the auto-spectra of the magnetic channels of the local station, which are assumed to be free of noise. This noise forms a smaller cluster in the left-hand side histograms (Fig. 2.19a).

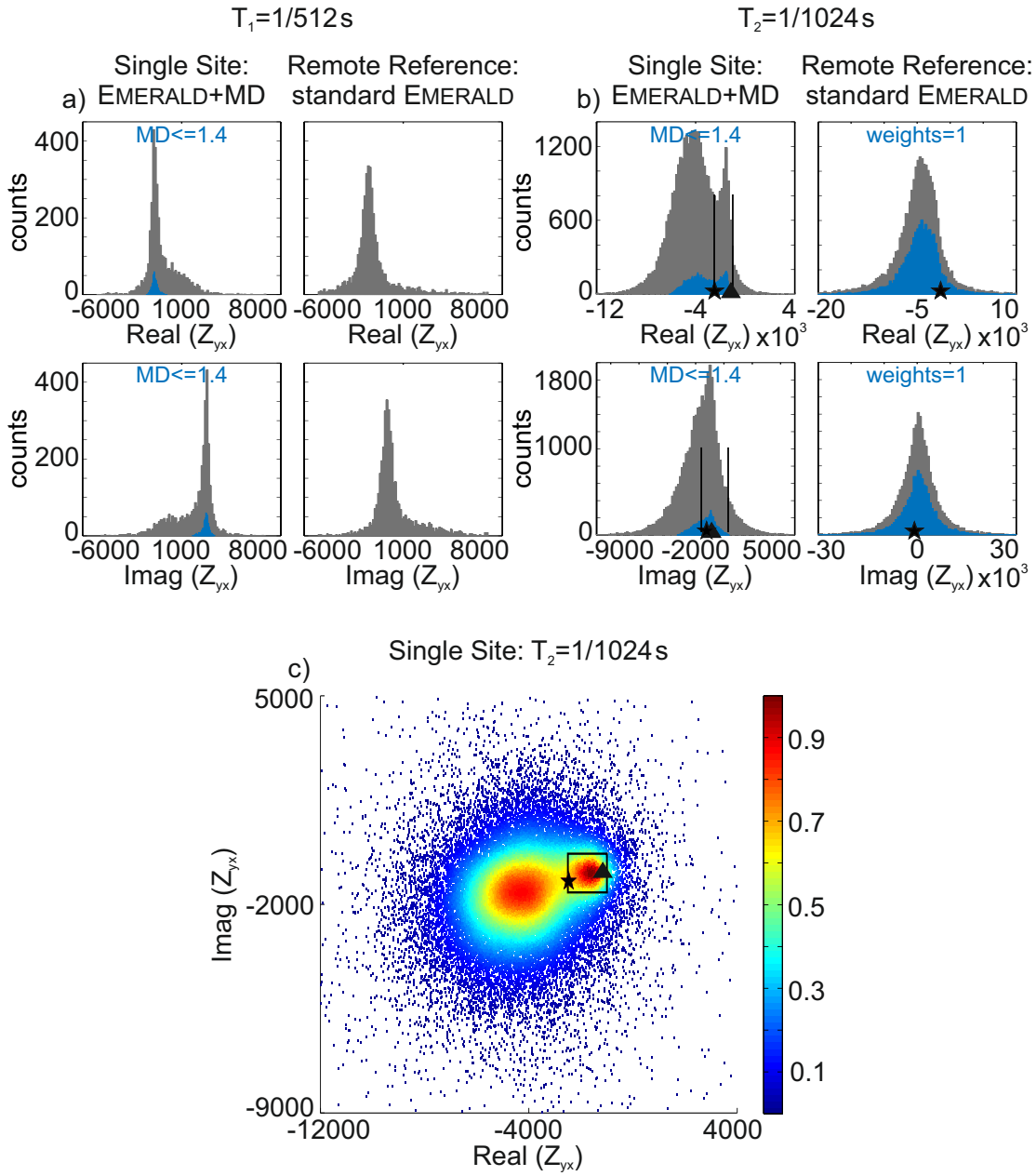

Figure 2.19: Histograms for a) $T_1 = 1/512\,s$, b) $T_2 = 1/1024\,s$ and c) colour-coded scatterplot for $T_2 = 1/1024\,s$ for Z_{yx} of station T-420. The black triangles represent the processing results of the undisturbed adjacent period of $T = 1/724\,s$ and the black asterisks indicate the actual processing result for the selected period. The distribution in the left-hand side in a) is tailed, whereby the tail can be removed by the MD criterion or the remote reference processing (right-hand side histograms). In b) two distributions are merged to a large extent also visible in the corresponding scatterplot in c). The black lines on the left-hand side represent the borders of the smaller cluster enclosed by the box in c) representing the desired MT signal.

If the amount of noise is relatively small, the MD criterion succeeds, shown by the blue-coloured distributions in Figure 2.19a (left-hand side). The remote reference processing also can remove this noise by replacing the auto-spectra by cross-spectra between the local and remote station (Fig. 2.19a right-hand side). However, if the amount of noise is relatively large, the MD criterion will fail. This is the case for the period of $T = 1/1024\,s$ (Figs. 2.19b & c), for which histograms and a scatterplot reveal a second large data cluster merged with the desired MT cluster. Processing results of the undisturbed adjacent period of $T = 1/724\,s$ are used to decide which cluster is related to the desired MT signal. Here, the desired signal belongs to the smaller cluster enclosed by the black box in Figure 2.19c consisting of the minority of all data. Consequently, the application of the MD criterion to the single site processing approach removes large parts of the true signal and will mainly focus on events originating from EM noise (Fig. 2.19b). Therefore, the transfer function estimation will be disturbed (second highlighted period in Fig. 2.18b). In contrast, the remote reference processing can successfully remove this noise by replacing the auto-spectra with cross-spectra resulting in only one data distribution (right-hand side of Fig. 2.19b).

2.5.6 Application to vertical and inter-station transfer functions

Although I focused mainly on impedances so far, the MD criterion can be used in the same way for estimation of vertical and inter-station transfer functions.

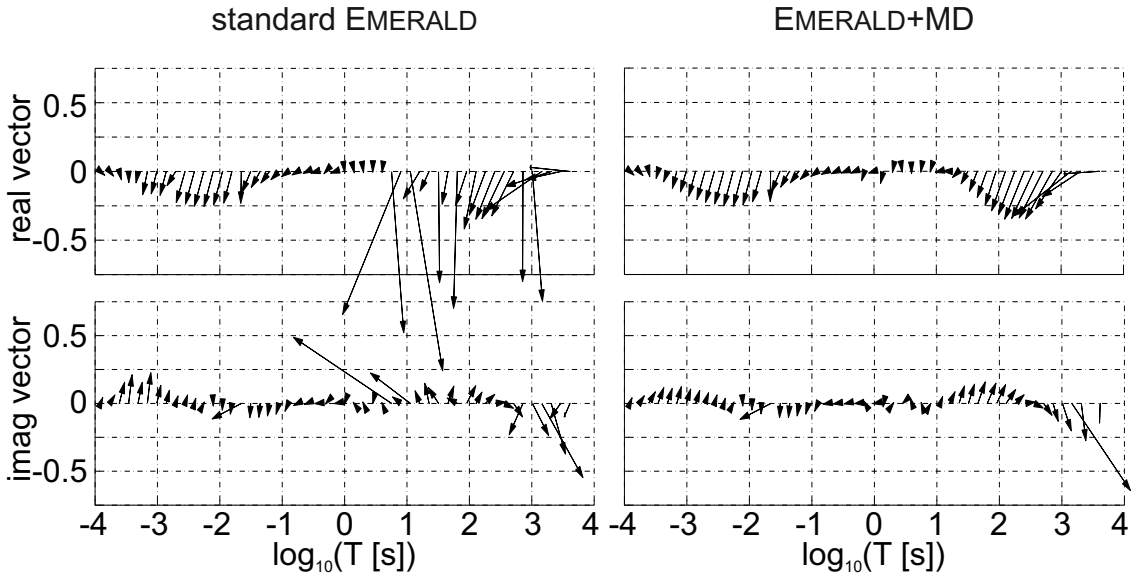

Figure 2.20: Real and imaginary induction vectors of station SA-208 comparing standard EMERALD and EMERALD+MD processing. The MD criterion improves the result leading to smoothly varying induction vectors with period.

To demonstrate this, improved vertical magnetic (Fig. 2.20) and inter-station transfer functions (Fig. 2.21) are shown for stations in South Africa. The induction vectors (Fig. 2.20) are almost completely smooth over the entire period range after the application of the MD criterion. Also the real and the imaginary parts of the inter-station transfer function components (Fig. 2.21) are much smoother for the EMERALD+MD processing in comparison to the standard EMERALD processing.

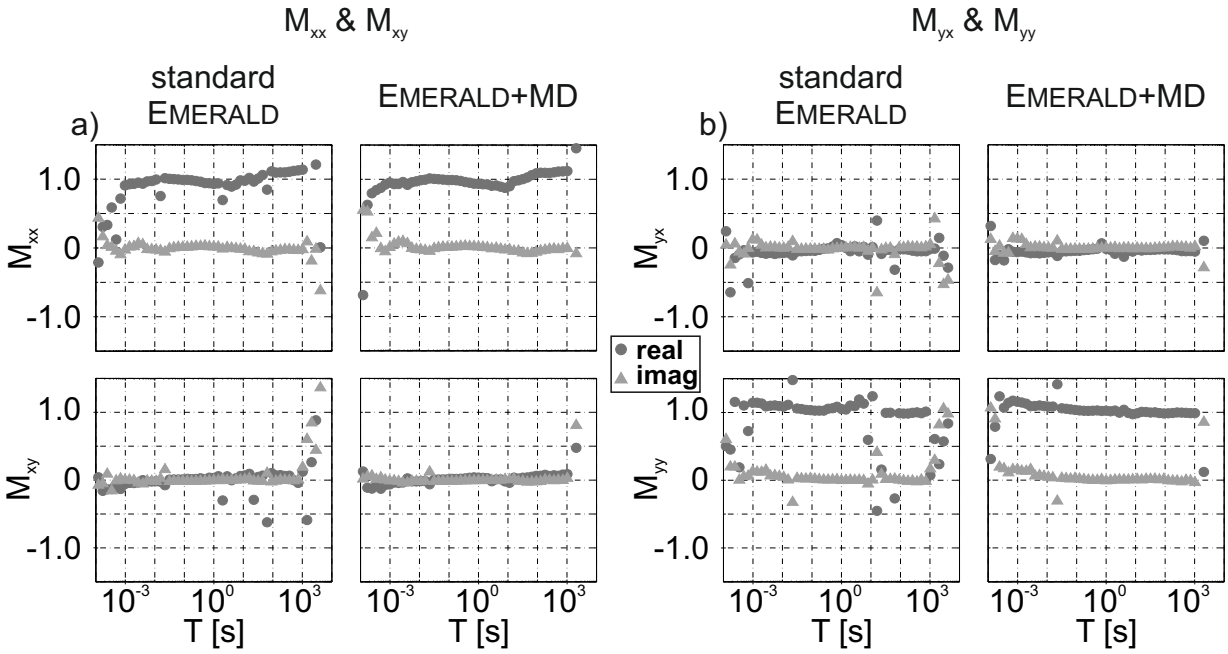

Figure 2.21: Inter-station transfer functions of station SA-220 with SA-518 comparing standard EMERALD and EMERALD+MD processing for a) M_{xx} and M_{xy} and b) M_{yx} and M_{yy}. The results of the MD criterion are superior and lead to smoother curves.

2.6 Conclusion of the Mahalanobis distance criterion

The MD criterion was implemented into the software package EMERALD to detect and remove outliers and noise. Consequently, this criterion confines the data to an ideally noise-free subset that is subsequently used in the stacking process. As input data, the real and imaginary parts of the transfer function components are used, as these quantities are the target quantities of the data processing. Single events that have a large distance to the estimated data centre under consideration of the covariance matrix are rejected from further processing. Events that are accepted by the MD criterion are used in the robust stacking algorithm.

The robust statistics within the standard EMERALD processing are able to remove large outliers in the tail of the distribution as well as a minor fraction of noise, but often fails if the data are affected

by a higher amount of noise. The MD criterion expands the robust statistics by a second statistical approach that confines the data to a subset. This subset ideally contains no outliers and significantly less noise affected data than the original data set. Therefore, the subsequent robust statistics within the stacking algorithm are able to calculate undisturbed MT transfer functions.

The new criterion was tested for several MT data sets from all over the world, which suffer from different noise contaminations. Comparison of standard EMERALD processing with EMERALD+MD processing reveals the circumstances under which the application of the MD criterion can improve the transfer function estimation. As the MD criterion is a statistical measure, it cannot distinguish between physically reasonable and non-physical data. Therefore, it fails for stations affected by a high amount of noise (higher than 50 %) or in cases where noise and desired MT clusters are superimposed to a large extent. In these cases, the application of the MD criterion can result in either totally misleading transfer functions by removing all or almost all desired MT signal or the EMERALD+MD processing does not show any improvements due to the large influence of noise. In such complicated cases, it is necessary to manually remove some noise, e.g. by further physically based data selection criteria or interactive selection algorithms, or to use other a-priori information to ensure that the majority of all single events is well-behaved and that the processing results in correct and undisturbed MT transfer functions.

However, in general data quality of stations with less than 50 % noise contamination can be significantly improved over the entire period range, even in the so-called dead band. EM noise often forms a completely independent cluster of transfer functions. As long as the majority of all data is well-behaved, the MD criterion is able to remove such clusters as well as to reduce scatter around the desired MT cluster by focusing on events close to the data centre. For these cases, the MD criterion is a useful measure to remove noise as well as outliers and therefore improve the MT transfer function estimation in an automated manner.

2.7 Introduction of the magnetic polarisation direction criterion

As illustrated in the last section, the MD criterion can and will fail if stations are affected by a high amount of noise, typically higher than 50 %. In these cases, some noise has to be removed manually, e.g. by physically data selection criteria or other a-priori information to ensure that the majority of all data is well-behaved.

Weckmann et al. (2005) proposed an interactive selection scheme that enables the selection of events based on visual inspection of several physical and statistical parameters, such as coherences or polarisation directions of the electric and magnetic wave field. This selection scheme is part of the EMERALD software package. Although in the past MT data of many stations have been improved, application is often tedious and time consuming. Furthermore, the result depends on the user's experiences.

I decided to use the magnetic polarisation (incidence) direction as an additional parameter to decide which cluster represents noise and subsequently remove them before the application of the MD criterion. In contrast to the electric field, the magnetic field is not assumed to show a preferred polarisation direction. The magnetic field is generated by a variety of different sources, e.g solar activity, ionospheric currents and lightning and thus the incidence directions of the generated magnetic fields vary in terms of their polarisation directions over 360°. For this reason, the magnetic polarisation direction is an appropriate parameter to detect and remove events originating from noise sources.

The magnetic polarisation direction angle α_B is calculated from the spectra for each event of a given target period:

$$\alpha_B = arctan \frac{2 * Real\left(\left[B_x B_y^*\right]\right)}{\left[B_x B_x^*\right] - \left[B_y B_y^*\right]}. \tag{2.12}$$

Although α_B theoretically can takes values between $0 - 360°$, we only calculate them in an interval of $[-90°, 90°]$. This is due to the utilisation of the simple "atan" function in C++, which cannot distinguish between different quadrants. The same interval is used in Weckmann et al. (2005) for the polarisation directions of the electric and magnetic field. The projection of all angles into only half of the complete range can be justified, because the magnetic polarisation directions of natural EM sources should vary without showing preferred directions. Therefore, preferred polarisation directions will only be linked to signals from EM noise. However, a possible and logical extension for the future would be to use the full range of all possible polarisation directions. The implementation of the magnetic polarisation direction criterion in EMERALD and its application is described in the next sections.

2.8 Implementation of the magnetic polarisation direction criterion

In comparison to the MD criterion, add-on of the magnetic polarisation direction (MPD) to the standard MD criterion turns this approach into a physically based one.

The MPD criterion can be divided into three main steps:

- In the first step, the magnetic polarisation direction angle α_B is calculated for each single event using equation (2.12).

- The second step consists of the kernel routine of the MPD criterion, the iterative selection algorithm. The basic idea is to find polarisation directions that exhibit an exceptionally large number of events and to remove the corresponding events. In order to decide which number of events is exceptionally large, the actual number of events is compared with a uniform distribution. Furthermore, the events are arranged in a consecutive order, which allows to remove only disturbed parts. Most of the following thresholds were found empirically after testing many stations with different polarisation patterns. In general, one can distinguish three different cases (see Fig. 2.22): a) undisturbed periods without a preferred polarisation direction, b) periods with preferred polarisation directions caused by EM noise and visible as "polarisation band" and c) periods suffering from more complex polarisation pattern.

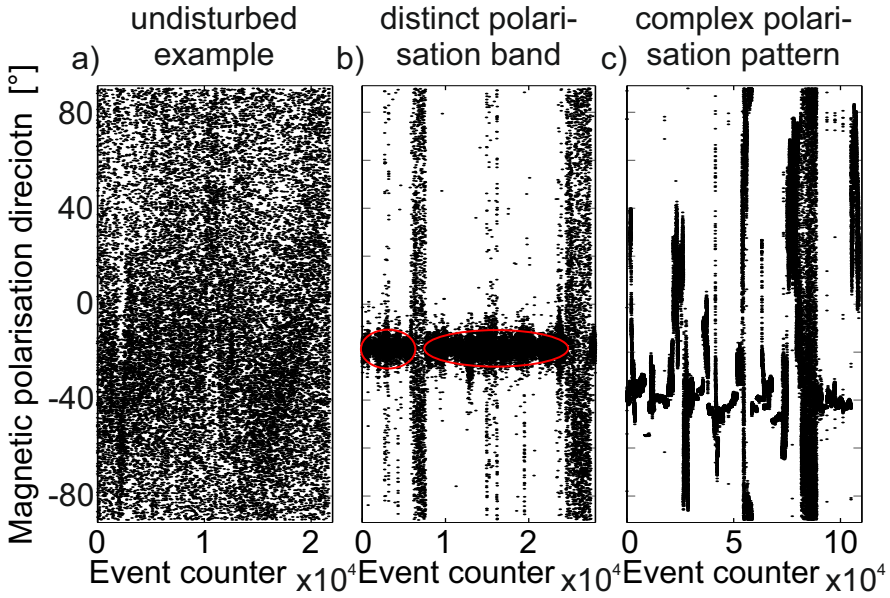

Figure 2.22: Plots of the magnetic polarisation direction angles of all events in an interval of $[-90°, 90°]$ for station V-304 for three different cases: a) undisturbed period, b) period temporarily suffering from magnetic polarised EM noise signal visible as distinct polarisation band (red ellipses) and c) period with complex polarisation pattern.

At the beginning, all angles α_B are organised into a histogram. For this histogram, I chose 180 classes, so that each class has a width of 1°. The expected value E_k is the same for each class as I

assume a uniform distribution:

$$E_k = \frac{Number\ of\ accepted\ events}{180}. \tag{2.13}$$

Initially, all accepted events of the previous data selection criteria are unflagged.

In general, statistics are only used if a minimum number of observations or events exist. In the standard EMERALD processing, the robust stacking algorithm is applied to all periods consisting of at least five events. Therefore, it makes sense to distinguish different cases depending on the amount of accepted/unflagged events and to treat these cases differently.

In the first case, I consider periods that only have a small amount of events compared to the number of classes. The MPD criterion distinguishes two different groups: (i) periods with less or equal 90 events and (ii) periods with less or equal 180 events representing expected values of 0.5 and 1, respectively. This particularly applies for the long periods. Due to the very limited number of available events, no statistics or iterative scheme are applied. Instead, the algorithm marks in one step all classes with an event number tenfold larger than these expected values.

The second case considers periods with less or equal 900 accepted/unflagged events (but > 180). Again, due to the limited number of events ($1 \leq E_k \leq 5$) no statistics are applied, but iterations are now allowed. In a first step, the algorithm searchs the class with the maximum number of events. In a second step, the actual number of events x_{max} in this class is compared to the expected value. If the difference between the actual number of events and the expected number is tenfold larger than the expected value (eq.2.14) , then all events in this class are flagged and the iterative algorithm is continued. Otherwise, the iterative algorithm stopps, because no preferred polarisation direction could be found.

$$\frac{x_{max} - E_k}{E_k} > 10 \tag{2.14}$$

The third and last case treats all periods with number of accepted/unflagged events larger than 900. For this group, the expected value for each class is at least 5 and therefore it is possible to apply statistics. Similar to the second case, the algorithm searchs in a first step for the class with the maximum number of events. If the number of events x_{max} in this class is exceptionally large, a statistical approach is started to remove only events belonging to disturbed segments (see e.g. highlighted areas in Fig. 2.22b). Thereafter, the iterative algorithm is continued. If a preferred polarisation direction is not observed, then the iterative algorithm stopps.

Tests have shown that the distribution of all polarisation angles get closer to a uniform distribution

with an increasing number of available events. Therefore, the limit of events in a class which is characterised as exceptionally large is lowered in a way that the difference between the actual number of events and the expected value only has to be three times as much as the expected value (instead of ten times as in the first two cases):

$$\frac{x_{max} - E_k}{E_k} > 3. \tag{2.15}$$

For the statistical approach to remove only disturbed segements, the class with the maximum number of events is marked and the total set of events is divided into sample windows of 300 events, respectively. Each of these windows is evaluated separately. As usual, first the algorithm calculates the number of events x_i that belong to the marked class for the $i-th$ window.
Secondly, the number of events x_i is compared to the maximum number of events x_{max} in the marked class:

$$x_i > 1.\bar{6} * x \tag{2.16}$$

with

$$x = \frac{x_{max} - E_k}{E_k}. \tag{2.17}$$

The factor $1.\bar{6}$ results from the ratio of window length divided by the amount of classes. If x_i is large (see eq. 2.16), the corresponding events in this window are flagged.

- In the last step, all flagged events are rejected.

The MPD criterion was intensively tested for many stations with different noise contaminations. The chosen limits were set after evaluating stations with no preferred magnetic polarisation direction as well as stations with complex polarisation pattern. These limits are selected in a conservative manner, so that the MPD criterion only removes events if there exists a definite polarisation direction. For stations that do not show any preferred polarisation direction, these conservative parameters ensure that all events are accepted. In contrast, for stations that are highly affected by polarised noise, this criterion sometimes does not remove all events which would be removed by visual inspection. The following examples in the next section will show the effectiveness of this approach. Only for complicated polarisation pattern and for highly noise contaminated stations the MPD criterion is too conservative. However, in these complicated cases it is always recommended to select distorted events by additional visual inspection.

The work flow of the MPD criterion is summarised in the appendix (Fig. A.3).

2.9 Application of the magnetic polarisation direction criterion

I will present two examples to illustrate which kind of polarisation patterns can be removed by the automatic MPD criterion. To compare the MPD criterion with the standard MD criterion, processing results of standard EMERALD and EMERALD+MD processing with and without application of the MPD criterion are compared as apparent resistivity and phase curves of one off-diagonal component. All processing examples use the coherence criterion with a threshold of 0.9. Furthermore, plots of the magnetic polarisation direction of all events are shown for selected periods.

2.9.1 Distinct polarisation bands

Standard EMERALD processing leads to poor MT results (Fig. 2.23a); especially the period range between 0.1 s and 1 s is heavily distorted. Application of the MPD criterion can almost completely improve the processing results in this period range; however, it cannot improve the short periods as the corresponding events do not exhibit a preferred magnetic polarisation direction (Fig. 2.23b). The EMERALD+MD criterion can improve the short periods, but fails to improve the period range, which suffers from magnetic polarised noise (Fig. 2.23c). Best results are obtained by a combination of both new criteria (Fig. 2.23d).

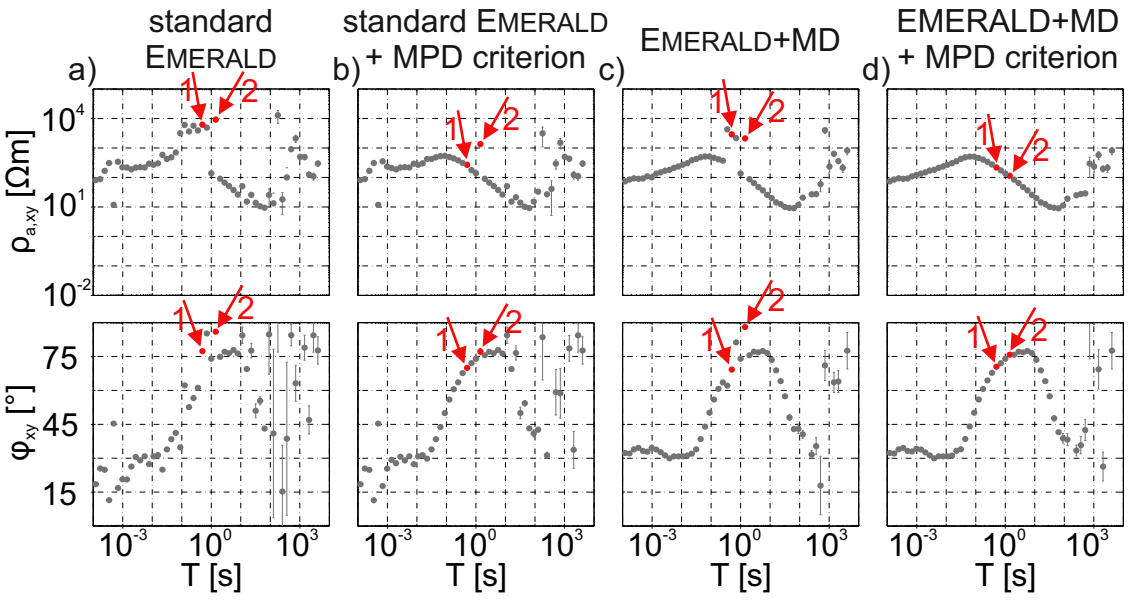

Figure 2.23: Apparent resistivity and phase curves of station SA-610 comparing the influence of the MPD criterion for standard EMERALD (a and b) and EMERALD+MD processing (c and d). The periods $T_1 = 0.5\,s$ and $T_2 = 1.4\,s$ are highlighted and further investigated in Figure 2.24.

For a better illustration of the polarised noise, two periods are selected for whose the distribution of all magnetic polarisation direction angles are shown before and after the application of the MPD criterion (Fig. 2.24). For the period of $T = 0.5\,s$ (Fig. 2.24a), a band of preferred polarisation directions can be observed between $-72°$ and $-64°$. This band is not continuous over all events, but is interrupted by some undistorted events. The application of the MPD criterion removes a large amount of events in the disturbed segments and results in around 50 % rejected events.

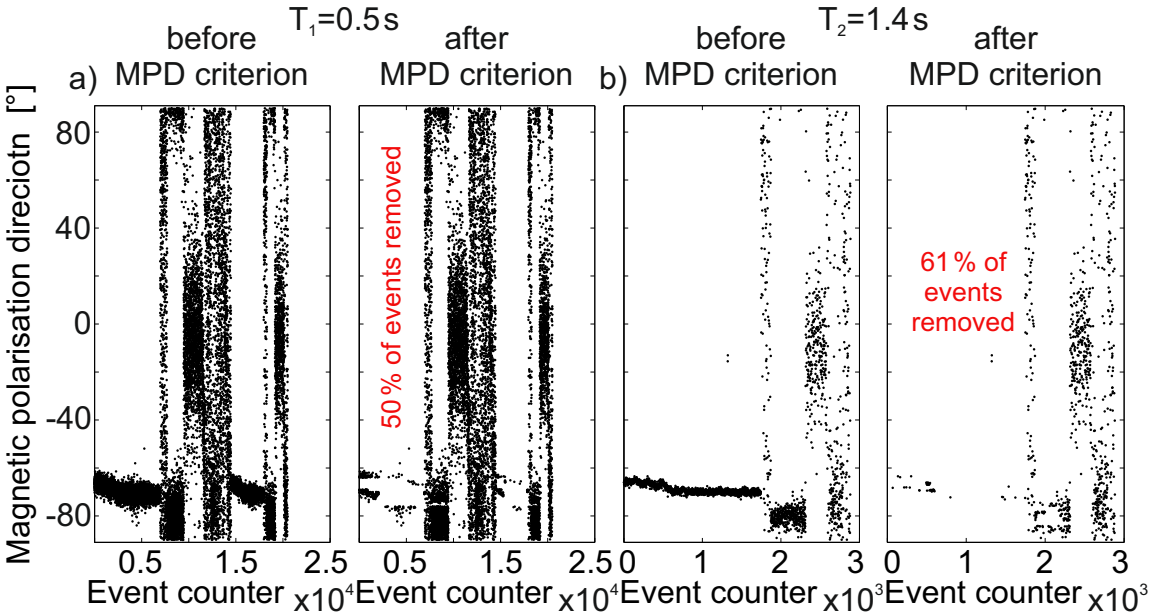

Figure 2.24: Plots of the magnetic polarisation direction angles of all events before and after the application of the MPD criterion for a) $T_1 = 0.5\,s$ and b) $T_2 = 1.4\,s$ in an interval of $[-90°, 90°]$. The events are displayed in consecutive manner. Most of the events belonging to a distinct polarisation band could be removed by the MPD criterion.

Due to the conservative implementation and the relatively large disturbed polarisation direction interval, the MPD criterion cannot remove all disturbed events. Some of them are falsely accepted, especially at the border of the selected interval. However, the MPD criterion removes enough bad events to improve the final processing result significantly. The EMERALD+MD processing fails for this period due to the high amount of noise, which is slightly above 50 %.

For the period of $T = 1.4\,s$ (Fig. 2.24b), two separate bands of preferred polarisation directions can be observed between $-72°$ and $-64°$ and between $-84°$ and $-76°$. Although both bands are almost completely removed, the MPD criterion is not able to produce enhanced results. Improved results are only observed in combination with the statistical MD criterion. For both periods, only disturbed parts of the distribution are removed as the total number of events is larger than 900. In these cases, the MPD criterion evaluates the entire distribution in small segments to ensure that only disturbed

parts are removed.

2.9.2 Complex polarisation pattern

Station V-304 is located in a deep sedimentary basin and should represent a homogeneous half-space.

Although the EMERALD+MD processing (Fig. 2.25c) can improve some parts of the curve in contrast to the standard EMERALD processing (Fig. 2.25a), it fails to revise periods, which are highly affected by polarised EM noise due to the large noise content exceeding 50 %. The MPD criterion can partly improve MT results for these periods (Fig. 2.25b). Best results are obtained, if the EMERALD+MD processing is combined by the MPD criterion (Fig. 2.25d).

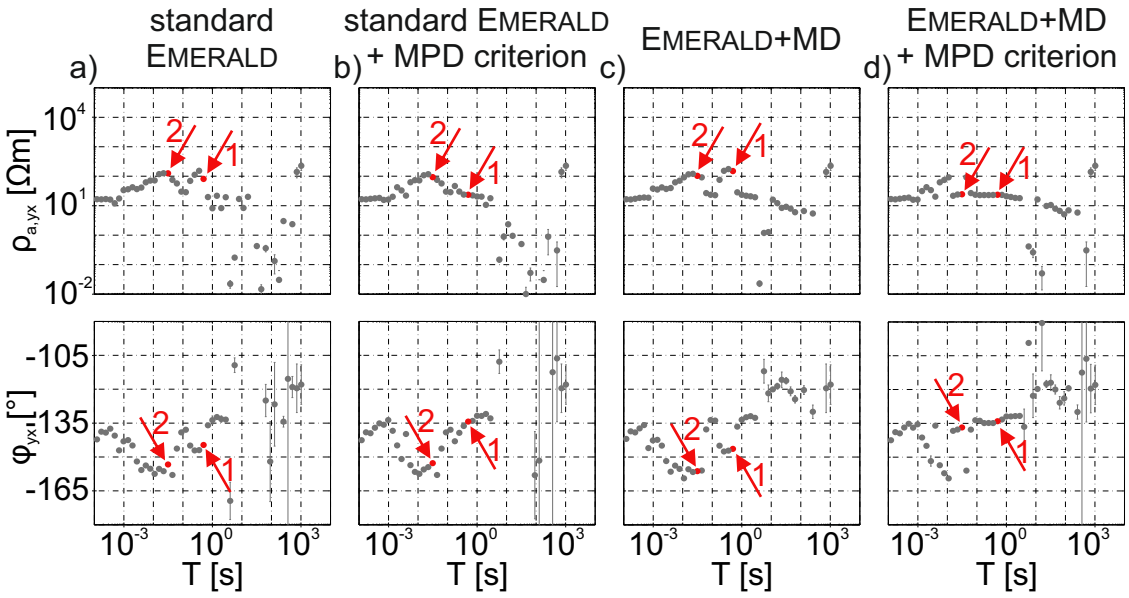

Figure 2.25: Apparent resistivity and phase curves of station V-304 comparing the influence of the MPD criterion for standard EMERALD (a and b) and EMERALD+MD processing (c and d). The periods $T_1 = 0.5\,s$ and $T_2 = 1/32\,s$ are highlighted and further investigated in Figure 2.26.

For a period of $T = 0.5\,s$ (Fig. 2.26a), the MPD criterion is able to improve the processing result on its own, because polarisation directions between $-27°$ and $-10°$ can be easily detected and removed by the MPD criterion. Due to the high amount of available undistorted events for this period, the algorithm can work with short segments to ensure that only disturbed parts are removed. In contrast, the period of $T = 1/32\,s$ suffers from a complicated polarisation pattern (Fig. 2.26b). The conservative implemented MPD criterion can only remove some parts of this pattern. Therefore, the application of the MPD criterion is not sufficient to improve the processing result. Only the combination of MD and MPD criterion is able to improve the result for this period.

Such complicated polarisation patterns are difficult to remove with an automatic criterion and require manual editing by an experienced user. The MPD criterion is only able to reject some parts of these patterns, which have an exceptionally large number of events. For more complicated structures, manual editing e.g. by interactive selection algorithms is highly recommended to obtain enhanced processing results.

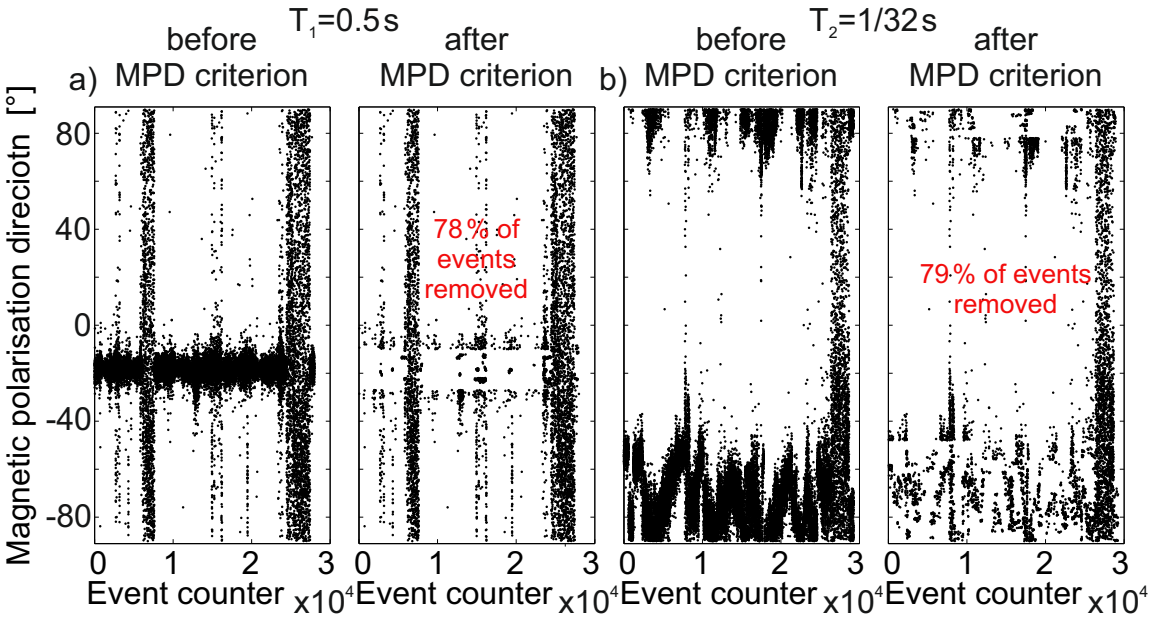

Figure 2.26: Plots of the magnetic polarisation direction angles of all events before and after the application of the MPD criterion for a) $T_1 = 0.5\,s$ and b) $T_2 = 1/32\,s$ in an interval of $[-90°, 90°]$. The events are displayed in consecutive manner. a) A broad polarisation band is visible and corresponding events are removed by the MPD criterion. b) A complex polarisation pattern dominates the majority of all events , which only can partly removed by the MPD criterion.

2.10 Conclusion of the magnetic polarisation direction criterion

I developed an automatic data selection criterion that recognises and removes events belonging to highly polarised magnetic signals. To evaluate the distribution of all polarisation angles, the distribution is organised in classes with a constant width. As a decision how many events are necessary to characterise a class as polarised, the actual number of events in a class is compared with an expected value assuming a uniform distribution. Only classes with an exceptionally large number of events in comparison to the expected value are marked and depending on the amount of events, further examined.

The parameters for the MPD criterion are chosen in a conservative manner after analysing many stations with different noise contaminations. The conservative selected parameters ensure that events are only removed for stations which exhibit a clear magnetic polarisation direction. Consequently, the MPD criterion is not able to completely detect and reject complex polarisation patterns. After the majority of distorted events is removed, the MD criterion can further improve the processing result. Otherwise, manual editing e.g. by interactive selection algorithms is essential to obtain enhanced processing results. However, for stations that suffer from noise with a preferred magnetic polarisation direction, the MPD criterion can significantly improve the processing results.

2.11 Chapter summary

I have developed two new data confinement and selection criteria and presented them in this chapter. While the MD criterion is statistically based, the add-on of the MPD brings in a physically based selection criterion. Both criteria can significantly improve the transfer function estimation.

MD criterion:

- A statistical MD criterion was developed to remove data points that have a large distance to the data centre under consideration of the data shape.
- For a robust MD calculation a deterministic MCD algorithm is used.
- The deterministic MCD algorithm was extended by a physically based initial estimator.
- The MD criterion fails if data have a large noise content ($> 50\,\%$) or in cases where transfer functions of EM noise overlap MT transfer functions to a large extent.
- For stations that are moderately affected by noise, application of the MD criterion leads to significantly improved results.

MPD criterion:

- The implemented MPD criterion removes events caused by a strongly polarised magnetic signal.
- This criterion is implemented in a conservative manner in order to ensure that only disturbed parts are removed.
- Consequently, it cannot completely reject events originating from complex polarisation patterns.
- For stations that suffer from noise with a dominant magnetic polarisation direction, application of the MPD criterion can significantly improve the transfer function estimation.

3 Magnetotelluric study of the Karoo Basin

Within the framework of the ongoing search for new energy resources, shale gas has become important as an alternative resource in the last years. In this context, the hydrocarbon bearing potential of the Karoo Basin in South Africa, with special focus on the black shales of the Whitehill Formation, has inspired the interest of the petroleum industry and scientists. The Karoo Basin, shown in Figure 3.1, covers large parts of southern Africa and represents about 100 million years of sedimentation.

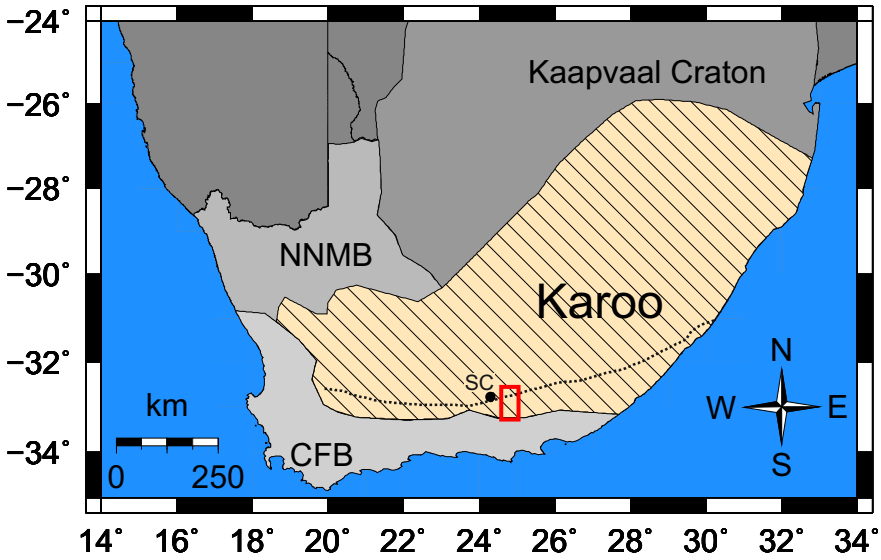

Figure 3.1: Simplified terrane map of southern Africa, modified after Weckmann et al. (2007a). The map shows the main tectonic units as the Archean Kaapvaal Craton, the Mesoproterozoic Namaqua Natal Mobile Belt (NNMB) and the upper Palaeozoic Cape Fold Belt (CFB). The Karoo Basin (shaded) covers large parts of southern Africa and is covered by Palaeozoic-Mesozoic sediments and igneous rocks. The maximum of the Beattie Magnetic Anomaly (BMA) is marked by a black dashed line. The location of the current study area is indicated by the red rectangle. The black dot labelled with SC is the location of the deep borehole SC3/67.

In 2013 the U.S. Energy Information Administration (EIA, 2013) published a study that estimates the worldwide shale gas resources. In this report, the shale gas reserve in South Africa was estimated as the eighth largest in the world. However, the true amount of potentially recoverable shale gas in the Karoo Basin is still unknown. In the last years, the estimated shale gas volume has strongly relativised due to different aspects such as thinning of the Karoo Supergroup to the north of the basin or thermal degassing of the shales by the intrusion of the Karoo dolerite suite.

In view of potential shale gas resources, an extensive research programme has been conducted, known as the “Karoo Shale Gas Baseline Research Programme”, with the participation of many national and

international scientific institutions. Main contributors are students and researchers from the African Earth Observatory Network (AEON) at Nelson Mandela University (NMU) in Port Elisabeth, South Africa. The aim of this baseline study is to obtain data and knowledge on key attributes and characteristics of the situation prior to the commencement of exploration and exploitation of shale gas. This programme summarises many different studies ranging from groundwater studies, structural geology, different geophysical methods, botanical and zoological subjects, to socio-economical applications. Many of these single studies are located in the area shown in Figure 3.1 due to the proximity of two shallow boreholes drilled by the NMU in 2012 (see e.g. Geel, 2013) and an electrical conductivity profile acquired by MT measurements in 2005 (Weckmann et al., 2007a) through a research collaboration between AEON and the German Research Centre for Geosciences (GFZ).

Previous MT studies (Weckmann et al., 2007a,b) along two different profiles in the Karoo Basin 350 km apart indicate that the target horizon for shale gas, the Whitehill Formation, can be mapped as a shallow horizontal band of high conductivity. This hypothesis is supported by a study of Branch et al. (2007). In this study the physical properties as well as the maturity of the carbon present in the Whitehill Formation were investigated by using core samples from boreholes. The results strongly suggest that the shallow conductive band in the MT models is linked with the carbon-rich sediments of the Whitehill Formation.

In the framework of this thesis, a MT field experiment was jointly conducted by GFZ and AEON in November 2014. There were two main goals of this experiment: (i) getting a background model of the electrical conductivity in this area to better define the geometry of the Whitehill Formation at depth and (ii) identifing shallow conductive features to indicate possible deep saline water reservoirs. One focus of this thesis was to improve data quality by developing and applying advanced processing techniques. Furthermore, the focus of this thesis was to develop a robust 3D background model of this region. A more detailed interpretation of this data set and the analysis of shallow conductive features have been done by J. Greve, a student from NMU (Greve, in preparation).

3.1 Geological background

The present extent of the Karoo Basin, as shown in Figure 3.1, is approximately $600,000\,km^2$ and it covers large parts of southern Africa (Cole, 1992; Geel, 2014). The basin represents about 100 million years of sedimentation and its estimated thickness ranges from $3\,km$ up to $12\,km$ with thinning of Karoo Supergroup to the north (Cole, 1992; Johnson et al., 1996; Svensen et al., 2007; Geel, 2014).

Many authors classify the Karoo Basin as a retro-arc foreland basin (e.g. Cole, 1992; Johnson et al., 1996; Catuneanu et al., 2005; Geel et al., 2013; Geel, 2014). The Karoo Basin is tectonically stable as it overlays the Namaqua Natal Mobile Belt in the south and the Archean Kaapvaal Craton in the north (Johnson et al., 2006; Geel, 2014). In the south, the basin is bounded by the Cape Fold Belt (Cole, 1992; Johnson et al., 2006). Approximately $200 - 300$ million years ago, the basin formed part of Gondwana. At that time, southwestern Gondwana was an assemblage of southern Africa, southern South America, East Antarctica, the microplates of West Antarctica and Falkland Islands. The Karoo Basin is part of a number of basins, which formed at that time between the uplifted landmasses to the south and the cratonic highland in the north (Geel, 2014). The tectonic regime was characterised by processes of subduction and orogenesis along the palaeo-Pacific (Panthalassan) margin of Gondwana (Catuneanu et al., 2005). Figure 3.2 illustrates the location and genesis of these basins.

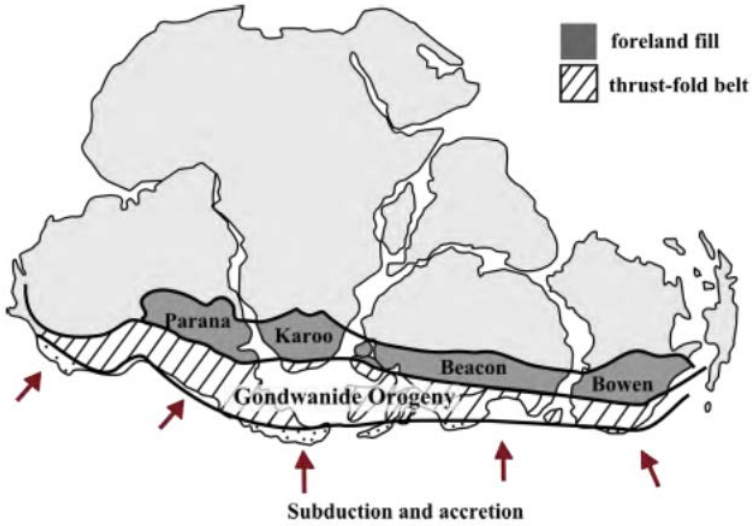

Figure 3.2: Schematic map showing the ongoing accretion tectonics along the southern margin of Gondwana during the late Palaeozoic ($300 \pm 75\,Ma$) from Geel (2014), modified after de Wit & Ransome (1992). The extensive foreland basins, such as the Karoo Basin, are shown in dark grey colours.

The compressional regime, associated with collision and terrane accretion, led to the formation of a $\sim 6000\,km$ long thrust-fold belt. The Cape Fold Belt in South Africa is a small portion of this belt (Catuneanu et al., 2005). It is suggested that the Karoo Basin evolved as a consequence of the formation of the Cape Fold Belt to the south during the shallow-angle subduction and accretion of the palaeo-Pacific plate beneath Gondwana during the late Carboniferous period (Cole, 1992; Catuneanu et al., 2005; Tinker et al., 2008; Geel et al., 2013).

The sediments of the Karoo Basin are divided into two supergroups, namely the Cape and the Karoo

Supergroup (Tinker et al., 2008; Geel, 2014). The Cape Supergroup is the older of the two groups and was formed prior to the Cape Fold Belt Orogeny 500 to 350 million years ago (Tinker et al., 2008; Geel, 2014). It is divided into the Table Mountain, Bokkeveld and Witteberg Group.

The Karoo Supergroup was deposited in the newly formed foreland basin related to the Cape Fold Belt Orogeny. Its sedimentation started during the late Carboniferous and the end of the Karoo sedimentation is dated to circa $185-180$ million years ago in the early Jurassic by the outpouring and intrusion of the Drakensberg basalts and dolerites (Thomas et al., 1993; Johnson et al., 1996; Tinker et al., 2008; Geel, 2014). The Cape Supergroup and at least the lower units of the Karoo Supergroup were intensely deformed along the southern basin edge due to the Cape Fold Belt Orogeny (Johnson et al., 2006).

During the Gondwana break-up 145 to 130 million years ago, the compressional Cape Fold Belt-aged orogenic structures were reactivated. The newly created accommodation space was filled with post-orogenic clastic sediments e.g. the Uitenhage Group, forming large basins in the south (Lock, 1978; Tinker et al., 2008).

The sediments of the Karoo Supergroup are of special economic importance as they contain e.g. extensive coal deposits and organic shales interesting for shale gas exploration (Johnson et al., 2006; Geel, 2014). Stratigraphically, the Karoo Supergroup is divided into several groups representing many contrasting sedimentological characteristics (Geel, 2014). From oldest to youngest the Karoo Supergroup comprises following groups: Dwyka Group, Ecca Group, Beaufort Group, Stromberg Group and Drakensberg Group (Geel, 2014), whereby only the first three groups can be found in the study area (see Fig. 3.4). These three groups as well as the formations containing black shales are summarised in Figure 3.3.

The Ecca Group is the second lowest formation of the Karoo Supergroup and contains 279 to 245 million years old black shales, whose shale gas bearing potential became of increasing interest to the petroleum industry recently. In the southern part of the Karoo Basin, the Ecca Group reaches a maximum thickness of $3000\,m$ (Catuneanu et al., 2005). During the time of the deposition of the Ecca Group sediments, melting of the extensive ice sheet resulted in a transgression and formed the marine Ecca basin (Johnson et al., 2006). Initially the basin was deep, and anoxic conditions allowed the suspension of fine grained materials and the preservation of organic material. The Prince Albert, Whitehill and Collingham Formations form the lower Ecca Group and were deposited during that time. There exists an abrupt change in depositional conditions between the Prince Albert and the Whitehill Formation. This change is attributed to the shallowing of the basin and termination of

oceanic circulation (Johnson et al., 2006). The Collingham Formation comprises turbidites and tuff beds expressing a change in tectonic conditions (Geel, 2014). Ongoing sedimentation and filling of the basin compacted fine grained organic rich sediment and led to the formation of black shales. The final filling of the aqueous basin took place with the deposition of the upper Ecca and the lower Beaufort Group. The inland sea was replaced by broad flood plains.

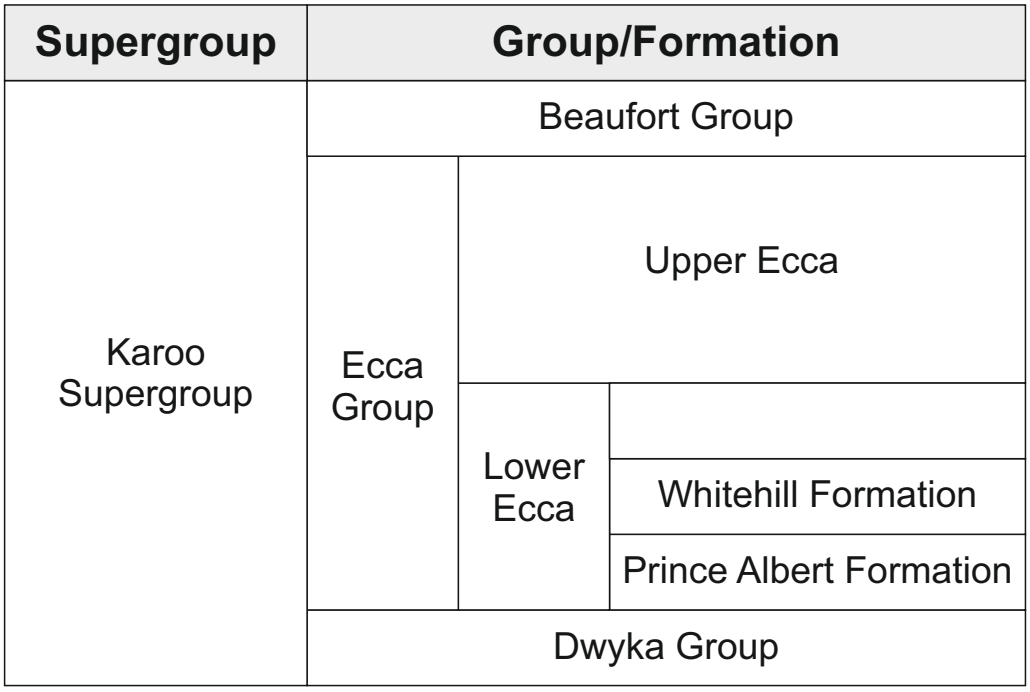

Figure 3.3: Simplified stratigraphy summarising the three groups of the Karoo Supergroup that exist in the study area as well as the two formations associated with the potential shale gas bearing layers.

The black shales of the Ecca Group are mainly focused in the Whitehill Formation (Johnson et al., 2006; Geel, 2014). The thickness of the Whitehill Formation varies from $10\,m$ to $80\,m$ in the Eastern Cape (Johnson et al., 2006; Geel, 2014). Tectonic as well as sedimentary environment provided the conditions for these deeply buried black shales to reach maturity levels in the gas window. The Whitehill Formation is known to contain large amounts of accessory pyrite (Geel, 2014). The black shales of the Whitehill Formation show many characteristics emphasing their shale gas bearing potential. However, the amount of exploitable gas is still unknown and is highly dependent on the influence of the Cape Fold Belt and the intrusion of the Karoo dolerite suite. Especially in the south, the Karoo sediments are within broad proximity of the deforming tectonic front of the Cape Fold Belt (Geel, 2014).

The shale gas bearing potential of black shales can be roughly estimated from their conductivity. Adao

(2015) investigated the dependence of the electrical conductivity of black shales for different factors. For this study, Adao (2015) used rock samples of Posidonia black shales from the Lower Saxony Basin. Main conclusion, with interest for my study, is that black shales consisting of immature organic carbon or having hydrocarbon-generating thermal maturities are not electrically conductive. The electrical conductivity for these black shales is mainly controlled by water content and porosity. The carbon content contributes only to a minor part to the conductivity. However, with higher thermal maturities, reaching e.g. meta-anthracite and graphite stage, the internal structure of the carbon can be rearranged leading to a significant increase of the electrical conductivity. Black shales having such high thermal maturities, e.g. by being exposed to very high temperatures, can be resolved as good conductors with MT, but it can be assumed that most of the gas has already been liberated.

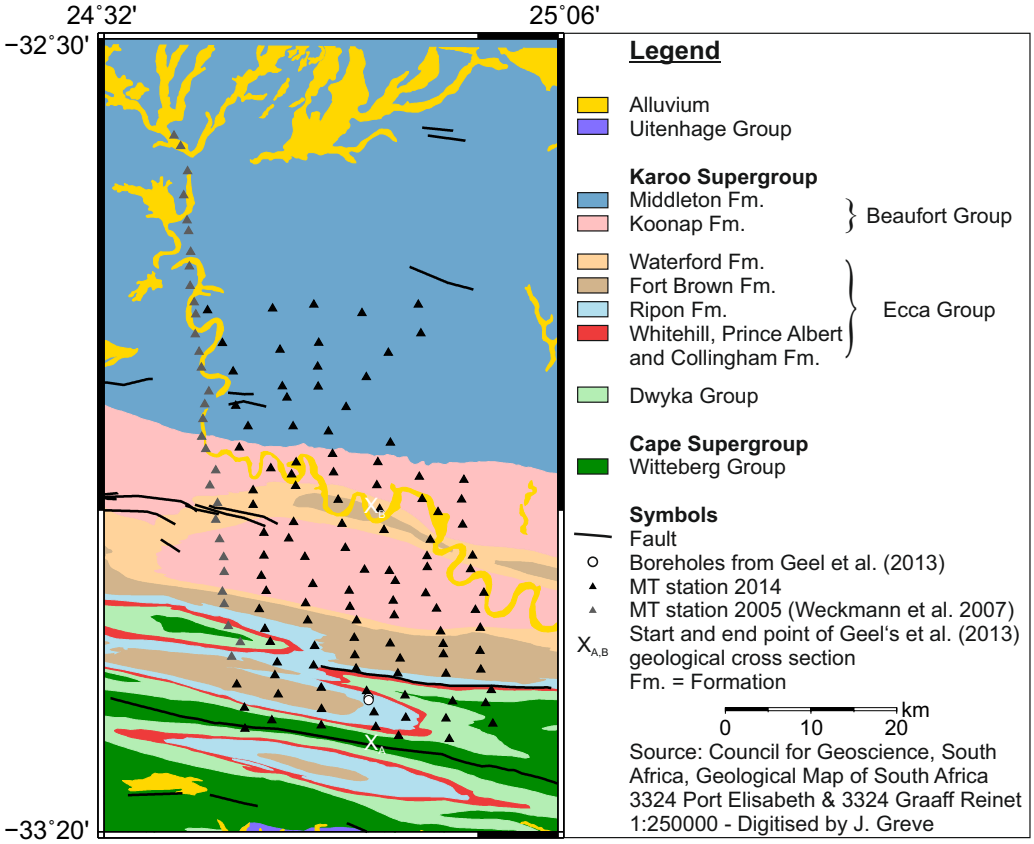

Figure 3.4: Simplified geological map of the study area. The location of MT stations of the two field experiments are marked by triangles. Two boreholes are located in the south and are represented by white circles. The Prince Albert, Whitehill and Collingham Formation are coloured in red.

The area of interest for the Karoo baseline study, shown in Figure 3.4, is underlain by 5 km to 6 km thick sediments of the Cape and Karoo Supergroup (Weckmann et al., 2007a). In the south, there are outcrops of the Witteberg Group as part of the older Cape Supergroup as well as deposits belonging to the younger Uitenhage Group. These outcrops are followed from south to north by deposits of the

Dwyka, the Ecca and the lower Beaufort Group. The Whitehill Formation in this area is estimated to be on average 30 m thick (Geel, 2014).

3.2 Existing scientific studies

Most of the existing scientific studies focus on geological issues. One of the latest and detailed geological study was conducted by Geel et al. (2013) and Geel (2014) in the Greystone area. The Greystone area is a small part of the current study area. The aim of this study was to obtain a comprehensive lithological, sedimentological, structural and geochemical description of the lower Ecca Group for a better understanding of e.g. its thermal maturity and gas bearing potential.

In the framework of Geel (2014), outcrops were mapped and the results were summarised to a detailed geological map and a south-north cross section, which is shown in Figure 3.5. The cross section illustrates the folded structure of the basin in the study area. While the Withehill Formation outcrops in the south, it deepens to the north and reaches a maximum estimated depth of 4 km.

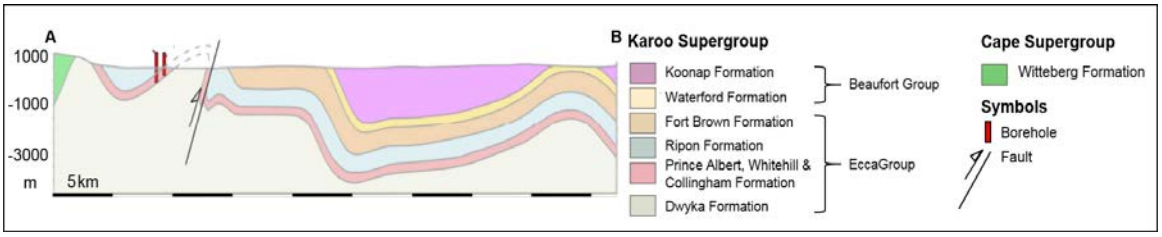

Figure 3.5: South-north cross section modified after Geel et al. (2013) and Geel (2014) of a small part of the current study area showing borehole locations together with the mapped fault in addition.

Two shallow boreholes, 100 m and 295 m deep, were drilled with the intention of extracting fresh core samples from the upper Dwyka Group, the Prince Albert, Whitehill and Collingham Formation as well as from the Ripon Formation. The location of the boreholes was chosen on the basis of the geological study assuming that the target formations are situated very shallow and partly outcrop, respectively. A thickness of $25-30\,m$ was observed in these two boreholes for the Whitehill Formation. This thickness corresponds well with the general assumed average thickness of this layer.

Geochemical and petrophysical analyses were applied to the core samples. The results show that the Whitehill Formation has a higher average total organic carbon value than the under- and overlaying Prince Albert and Collingham Formation. Furthermore, these analyses indicate that the lower Ecca Group rocks are overmature and that gas has been already liberated from the shale. The deep burial

of the Ecca Group sediments increased the thermal maturity in the past and formed the base of the development of gas. However, Geel (2014) assumed that an additional thermal maturation was developed in this area due to the proximity to the Cape Fold Belt. The determined high maturity suggests that the black shales can be highly electrically conductive. However, most of the core samples from these shallow depths were strongly weathered. This could reduce the electrical conductivity significantly. In general, the core samples from these shallow boreholes are not representative for my study as my target depth is much deeper. This can have a large influence on the electrical conductivity of the black shales.

Apart from geological studies, some geophysical studies were conducted in the past. The MT study by Weckmann et al. (2007a) was one of two MT studies in South Africa to resolve, among other things, the structural details of the Beattie Magnetic Anomaly (BMA). For this purpose, Weckmann et al. (2007a) recorded broadband MT data at 31 stations along a south-north profile in the vicinity of the town Jansenville in the Eastern Cape crossing the surface trace of the BMA. The MT data were recorded in a period range from $0.001 - 1000\,s$ with GPS synchronised S.P.A.M. MkIII and CASTLE broadband instruments. Magnetic and electric field components were measured with Metronix MFS05/06 induction coil magnetometers and non-polarisable Ag/AgCl telluric electrodes. A second MT study was conducted $350\,km$ farther west along a $150\,km$ long profile by Weckmann et al. (2007b). The BMA is one of the Earth's largest known continental geophysical anomaly. It extends for almost $1000\,km$ in east-west direction across the southern part of South Africa. In Figure 3.1 the maximum of this positive magnetic anomaly is marked by a dashed line. Although this anomaly was already discovered by Beattie (1909) a century ago, the source of this anomaly is still under debate. Both studies resolved a subvertical conductive anomaly beneath the surface trace of the BMA. The source of this narrow high electrical conductivity zone was interpreted by Weckmann et al. (2007a) and Weckmann et al. (2007b) as a mineralised shear zone, which cuts through a magnetic source responsable for the BMA. As the current study extends the study from Weckmann et al. (2007a) to the south and the east, the surface trace of the maximum of the BMA runs across the northern part of the study area. However, mapping of this anomaly was not the scope of the current MT study.

A second prominent feature in both previous MT models is a shallow, regionally continuous subhorizontal layer of high conductivities at $3\,km$ to $10\,km$ depth. This conductive layer is related to the Whitehill Formation. A study of Branch et al. (2007), discussed in the next paragraphs, supports this hypothesis. From geological studies, it is known that the Whitehill Formation has only a maximum thickness of several tens of metres. However, in MT models the thickness of a shallow conductive layer is typically poorly constrained as it is sensitive to the conductance (product of a layer's thick-

ness and its conductivity). Therefore, a layer could be thinner but more conductive or thicker but more resistive.

Branch et al. (2007) used samples from borehole SA1/66 to investigate the maturity of the carbon present in the Whitehill Formation as well as its physical properties. The borehole is located $50\,km$ west of the MT study from Weckmann et al. (2007b) and the Whitehill Formation is found in a depth range from $2750 - 2800\,m$. This depth range correlates with the shallow subhorizontal conductor in this MT study.

With impedance spectroscopy, Branch et al. (2007) could show that only samples from the Prince Albert and the Whitehill Formation showed high conductivities, whereby the conductivity of the Prince Albert Formation could be triggered by the high pyrite content in the used core sample. Since both layers contain black shales, one important property to make a link from electrical conductivity to shale gas properties is to look at the maturity. Vitrinite reflectance analysis is a standard tool to investigate the thermal maturity of organic matter. Based on this analysis the organic matter of the two formations were classified in the meta-anthracite maturity field. This is close enough to the graphite field so that the authors assume that the organic matter was responsible for the high conductivities.

Other resistivity studies of the Karoo Basin were carried out by the Council for Scientific and Industrial Research in the period of 1966 to 1980 (van Zijl, 1978; de Beer & Meyer, 1983, 1984; van Zijl, 2006). In the framework of these studies, more than 280 deep electrical soundings were conducted in different areas spread over the entire Karoo Basin. In areas where dolerite intrusions are absent, the investigation depth is up to $8\,km$. One of these areas (Graaff Reinet-Jansenville) overlaps with the study area of this thesis. 25 soundings were recorded in this area and the results were correlated with information of borehole SC3/67. This deep borehole, shown in Figure 3.1, is located west of the current study area. The top of the Whitehill Formation was estimated in $4130\,m$, which is close to the actual measured value of $3970\,m$ in the adjacent borehole SC3/67. The average resistivity of the Whitehill Formation obtained from the interpretation of all deep electrical sounding curves is $1\,\Omega m$ and the range is specified to $0.01 - 10\,\Omega m$. Due to the depth range of the Whitehill Formation, these resistivity values are indicative for the current study. The estimated average resistivity of the entire Graaff Reinet-Jansenville area is $700\,\Omega m$ with a range of $500 - 1100\,\Omega m$. In general, van Zijl (2006) observed an increase of the average sediment resistivity above the Whitehill with depth, probably caused by weathering effects, and laterally in a southerly direction as a result of diagenesis and low-grade metamorphism.

Furthermore, Stankiewicz et al. (2007) reported on a 200 km long wide-angle refraction seismic line running through the study area to perform a detailed analysis of the crust. As their study focused more on deriving a tectonic model for this region, small-scale structures were not resolved.

Further studies from the Karoo Basin within the Karoo Shale Gas Baseline Research Programme such as hydrogeology and groundwater monitoring, airbone geophysics and microseismics were already conducted, but analysis and interpretation is still ongoing.

3.3 Magnetotelluric measurements

The MT data considered in this thesis originate from two field experiments. In 2005, Weckmann et al. (2007a) recorded broadband MT data at 31 stations in the Eastern Cape. In the framework of this thesis, only the final impedance tensor data and VTFs from Weckmann et al. (2007a) were used. The data were not reprocessed. In November 2014, additional 111 stations were installed in the same region by colleagues from GFZ and from NMU in the framework of the Karoo Shale Gas Baseline Progamme.

3.3.1 Data aquisition

In 2014, 111 broadband five channel MT stations were installed covering an area of $30\,km \times 50\,km$. All stations recorded data in a period range of $0.0001 - 1000\,s$ using GPS synchronised S.P.A.M. MkIV instruments. Most of the stations recorded data for three days and at least three stations were operated simultaneously. Only station 312 were conducted almost during the entire field experiment as a local reference station. Magnetic and electric field components were, similar to 2005, measured with Metronix MFS05/06 induction coil magnetometers and non-polarisable Ag/AgCl telluric electrodes (from the Geophysical Instrumental Pool Potsdam). The newly measured data were aligned along six profiles parallel to the old profile from 2005. Three stations were added to the old profile in the south. The distance between two stations along a profile is about $2 - 2.5\,km$ and the distance between two profiles is roughly $5\,km$. Figure 3.6 displays the profiles with the MT stations.

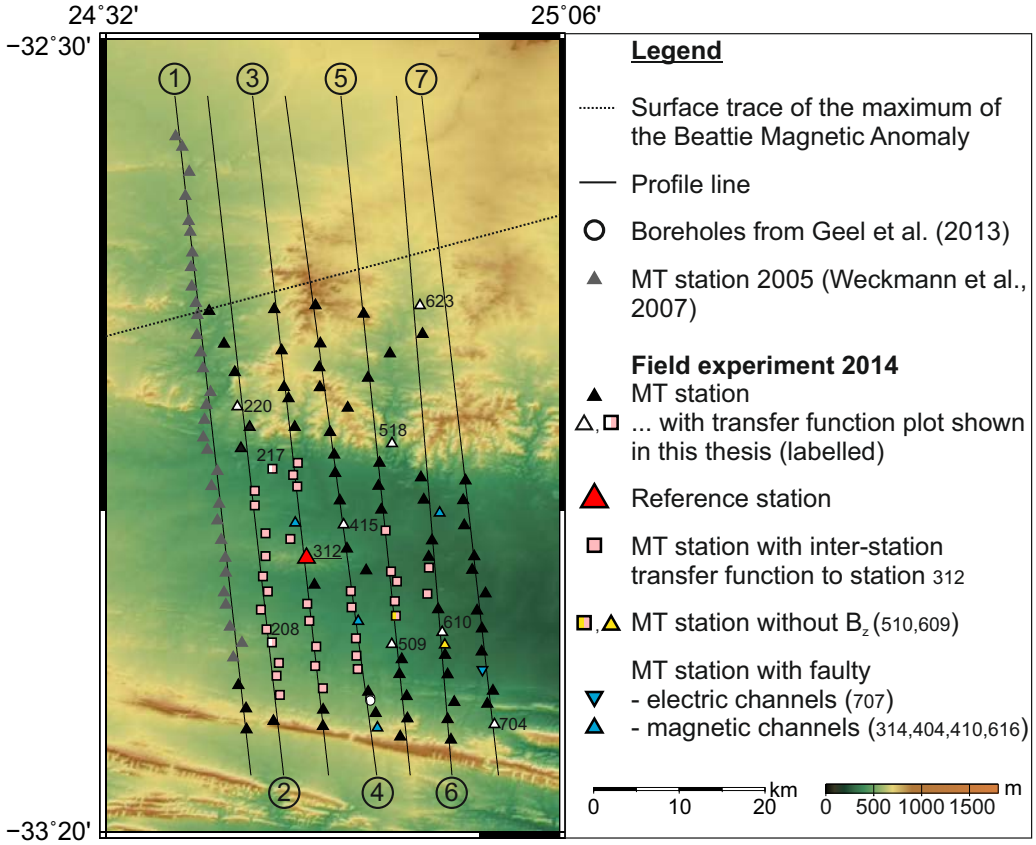

Figure 3.6: Map of study area in the Eastern Cape (South Africa), with locations of MT stations. Stations labelled with their number are shown as apparent resistivity and phase curves or induction vectors in this thesis.

3.3.2 Processing results

The two newly developed criteria, the MD and the MPD criterion, were applied to the entire data set from 2014 in addition with a coherence threshold. In general, this led to improved transfer function estimates as exemplarily shown in the previous chapter.

Pseudosections for the off-diagonal impedance tensor components as well as pseudosections for the VTFs of all seven profiles for (i) standard Emerald and (ii) advanced Emerald processing with MD and MPD criterion can be found in the appendix (A.4-A.10). Here, only an exemplary comparison between Emerald and the advanced Emerald processing is discussed (Fig. 3.7). The standard Emerald processing (Fig. 3.7a) results in many spurious data points, which do not agree with the assumption of smoothly varying transfer functions. These points have to be removed before an inversion that leads to less usable periods and subsequently to less information of the electrical conductivity structure. This affects especially the short periods ($T < 10^{-2}\,s$) disturbed by the $^1/_{50}\,s$ signal from power grids and its harmonics. But also longer periods ($T > 10\,s$) are affected, although

especially these periods are needed to resolve the target formations in greater depths. The application of the novel criteria leads to smoother transfer functions and reduces the number of poorly estimated data points (Fig. 3.7b), especially for short periods ($T < 10^{-1}\,s$) and for the long periods ($T > 10\,s$). Thus, we have more data points over a broader period range, which can be used in an inversion.

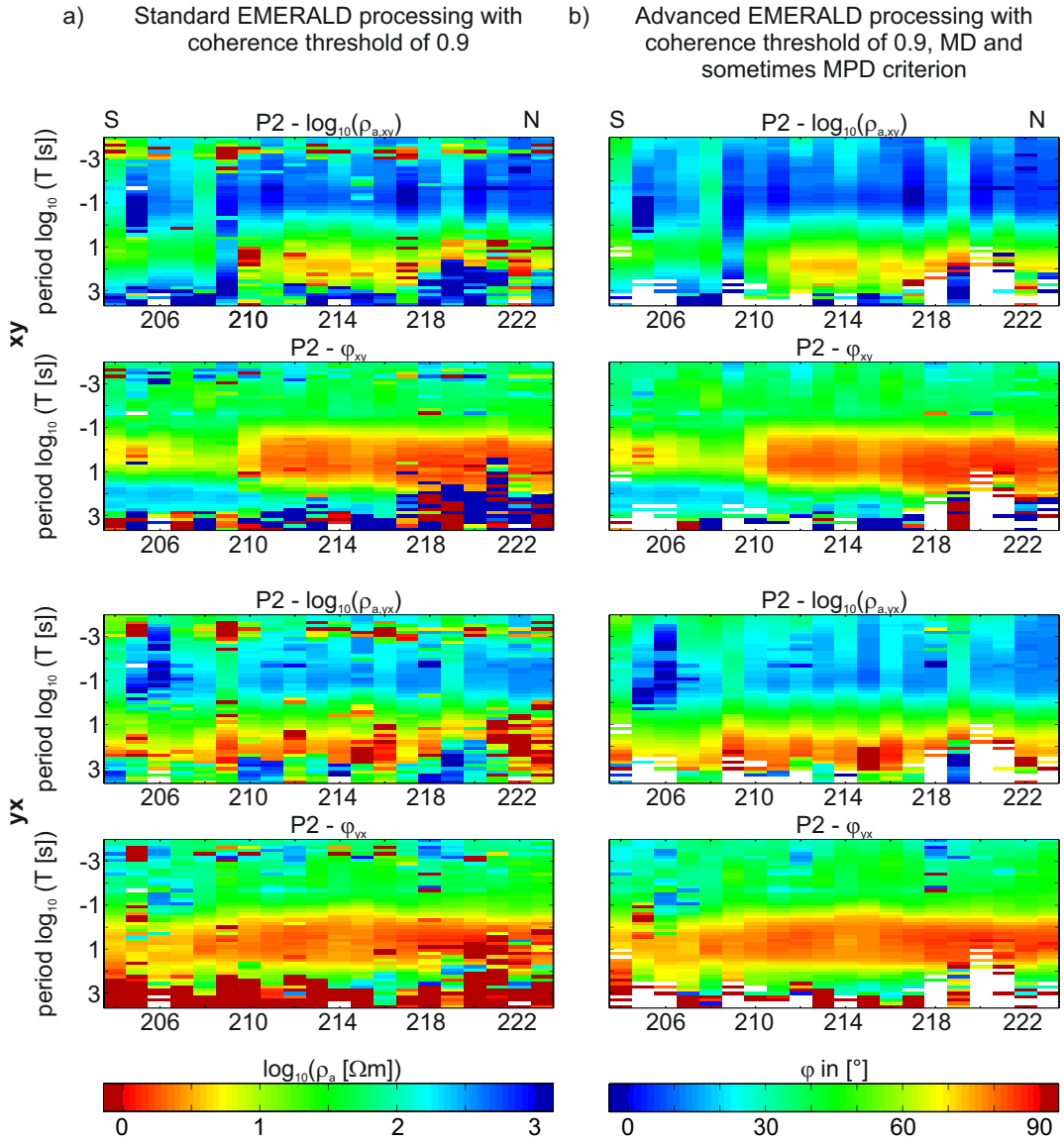

Figure 3.7: Comparison of standard EMERALD processing and advanced processing with additional application of MD and MPD criterion. In a) the apparent resistivity and phase values of both off-diagonal impedance tensor components are shown for all stations belonging to profile 2 along a pseudosection. The $y-axis$ of each plot represents periods and the data of the different stations are visualised in separate columns. In b) similar plots are shown for the advanced processing using the MD criterion and for some stations the MPD criterion in addition to the processing shown in a). White space represents not existing data.

In the following, full impedance tensor data (as apparent resistivity and phase) as well as VTFs (as induction vectors) are shown for three stations representative for larger areas (Fig. 3.8).

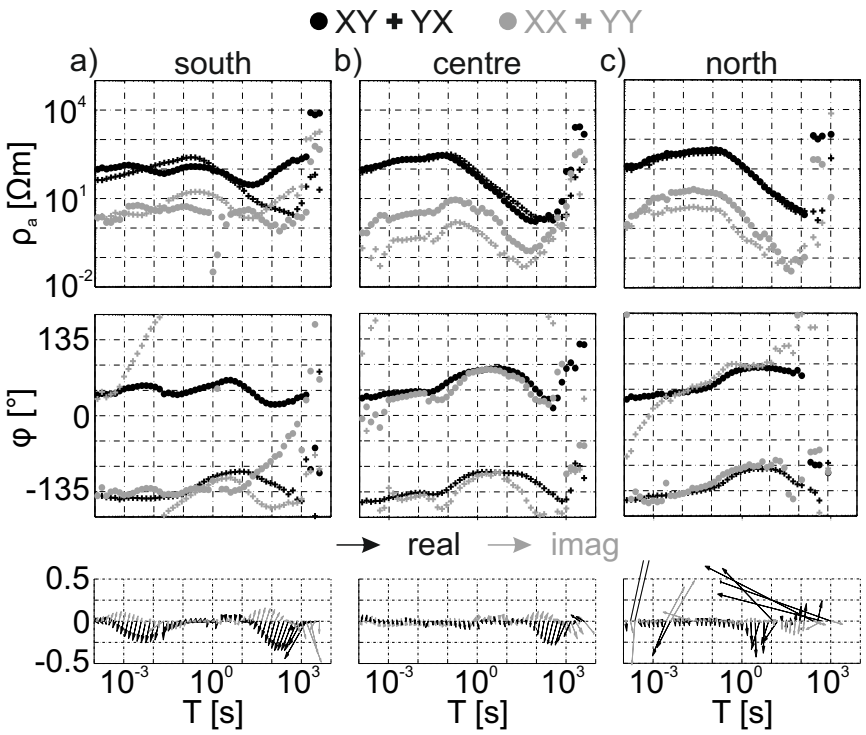

Figure 3.8: Full impedance data displayed as apparent resistivity (upper row) and phase (middle row) as well as induction vectors in Wiese convention (lower row) of three representative sites in a geographic coordinate system ($x \hat{=}$ geographic north, $y \hat{=}$ geographic east). a) Site 208 represents stations in the south of the study area. b) The reference site 312 is representative for the middle part and c) site 623 is located in the northern part of the study area.

The impedance tensor data are shown in a geographic coordinate system ($x \hat{=}$ north, $y \hat{=}$ east). All exemplary apparent resistivity and phase curves vary smoothly with period and are consistent.

For station 208, the general trend of the apparent resistivity curves indicate a change from resistive to more conductive to again more resistive structures with increasing period (Fig. 3.8a). The off-diagonal impedance components of this site show a large split for longer periods ($T > 10\,s$) and the main-diagonal impedance elements are at least one magnitude smaller than the off-diagonal components. The real and imaginary induction vectors are presented in Wiese convention (Wiese, 1962), so that the real induction vectors point away from good conductors. Induction vectors are large for the short ($T < 10^{-1}\,s$) and the long ($T > 10^{1}\,s$) periods with real vectors pointing in southwesterly direction and imaginary vectors pointing to northeast. Between this interval, the induction vectors become close to zero.

Site 312 and 623 show a similar behaviour for their impedance data (Figs. 3.8b & c). Both off-diagonal components have almost the same magnitude over the entire period range and indicate the same trend as station 208. However, for station 312 most of the induction vectors are small ($T < 10^{2}\,s$); for longer

periods the real induction vectors point southwest. The induction vectors for station 623 have a poor data quality representative for most of the induction vectors in the nothern part of the study area.

Processing results for the off-diagonal impedance tensor components for all stations are summarised in Figures 3.9 and 3.10. In addition, results of the VTFs are shown in the appendix (A.4-A.10).

3.3.3 Data interpretation

First conclusions of the electrical conductivity structure of the subsurface can already be drawn from the representative apparent resistivity and phase curves in Figure 3.8. As the curves of the middle and northern part look very similar and the main-diagonal components are relatively small, this indicates a layered (1D) electrical conductivity structure for large parts of the study area. However, the subsurface structure seems to become more complex in the south.

A more comprehensive overview can be derived by looking at pseudosections of all stations (Figs. 3.9 & 3.10).

The data in pseudosections are plotted against periods. The period axis indirectly represents a measure of depth as both penetration depths and induction volumes increase with longer periods. The data of different stations are plotted along a profile line. This allows a first estimate how the electrical conductivity structure changes along this profile direction. Furthermore, a comparison of different adjacent profiles enables to detect changes in another direction. All profiles in this study run roughly in south-north direction and are more or less parallel to each other.

At first sight, apparent resistivity and phase values look very similar for all profiles and for both off-diagonal impedance tensor components. They confirm the general trend from resistive to more conductive to again more resistive structures with increasing period, already indicated by the three exemplary stations (Fig. 3.8). A prominent high conductivity feature is indicated at greater depth. This is supported by high phase values for longer periods. For the longest periods phase values start to decrease again, while apparent resistivities remain low. This is a well-known and often observed behaviour that phases seem to sense deeper as they first indicate conductivity changes for a larger induction volume.

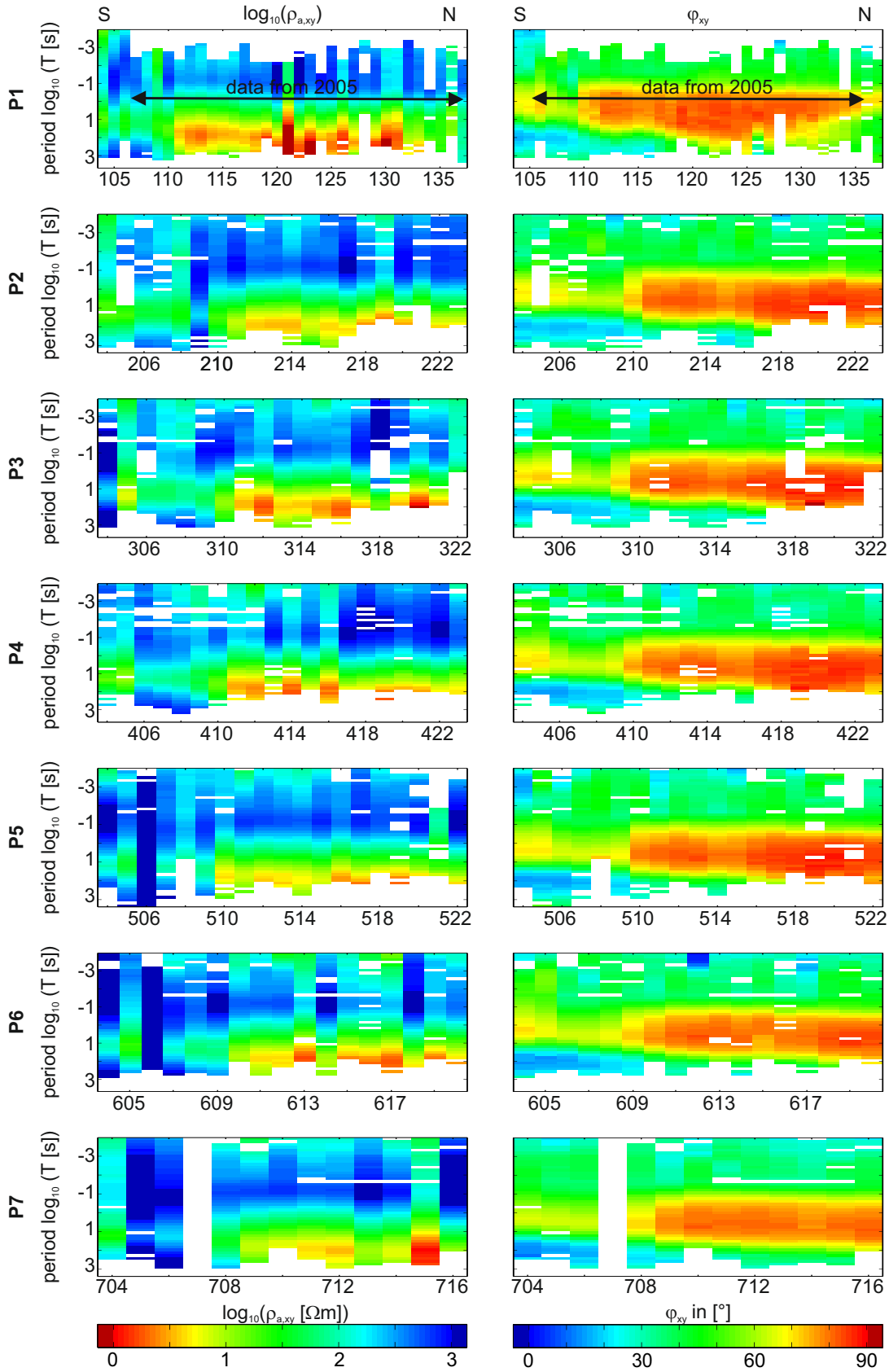

Figure 3.9: Masked processing results of Z_{xy} for all seven profiles as pseudosections of apparent resistivity and phase. For profile 1, only the first three stations were processed in the framework of this thesis as the other stations were measured and evaluated by Weckmann et al. (2007a). White space represents not existing or masked data.

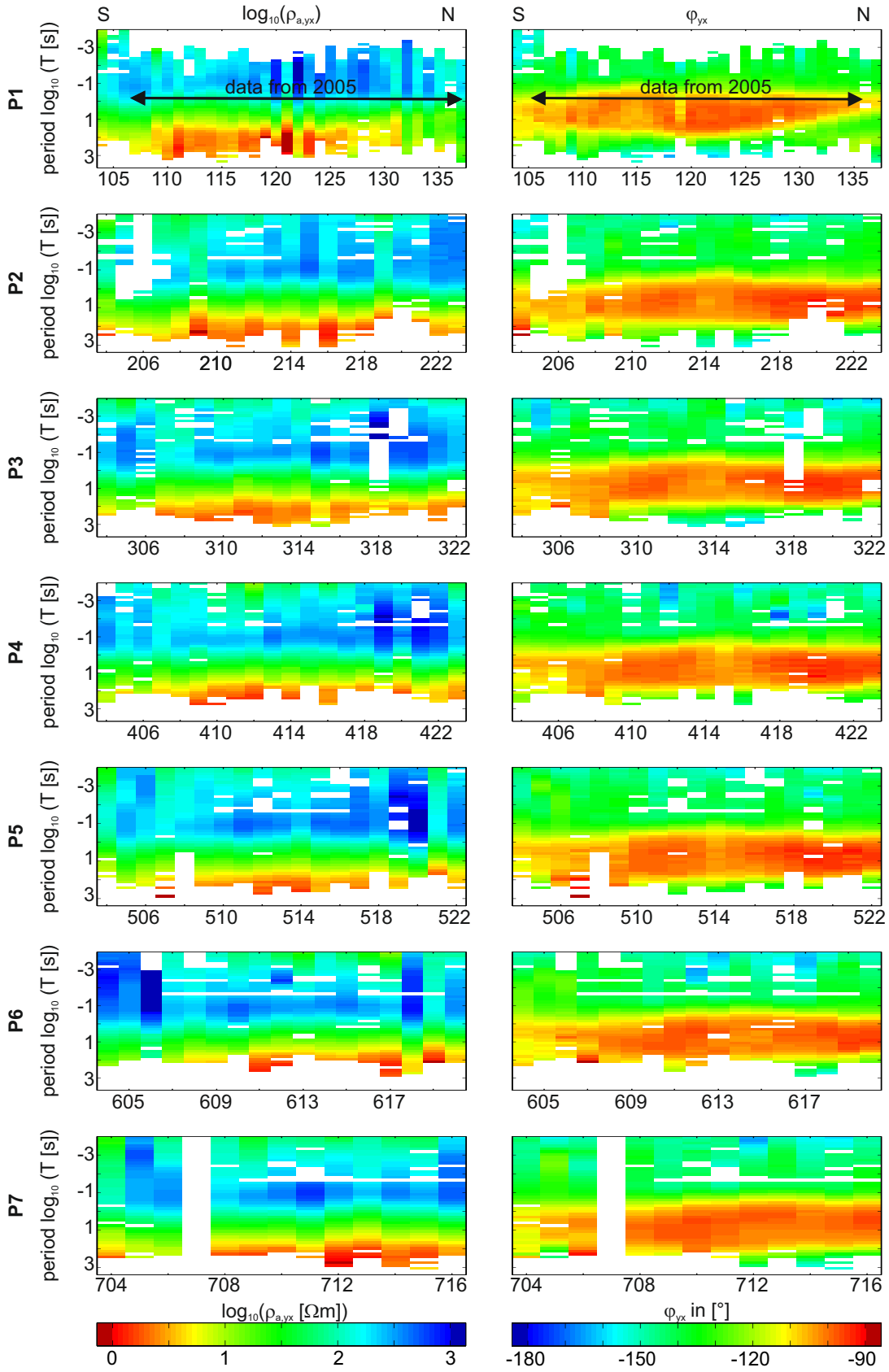

Figure 3.10: Masked processing results of Z_{yx} for all seven profiles as pseudosections of apparent resistivity and phase. For profile 1, only the first three stations were processed in the framework of this thesis as the other stations were measured and evaluated by Weckmann et al. (2007a). White space represents not existing or masked data.

However, a closer look reveals some differences. The most important difference is visible between Z_{xy} and Z_{yx} at the long periods ($T > 10\,s$). The apparent resistivity and phase pseudosections of Z_{yx} reveal a continuous band of high electrical conductivities from south to north. A similar electrically conductive band can be observed for Z_{xy}, but only in the central and northern part of the study area. This implies a more complex electrical conductivity structure in the south. Furthermore, the pseudosections suggest that the conductivity structure in the study area has more significant changes in south-north than in east-west direction.

Apart from pseudosections of impedances, maps of induction vectors (Fig. 3.11) can be used for an initial assessment of lateral changes of the electrical conductivity structure.

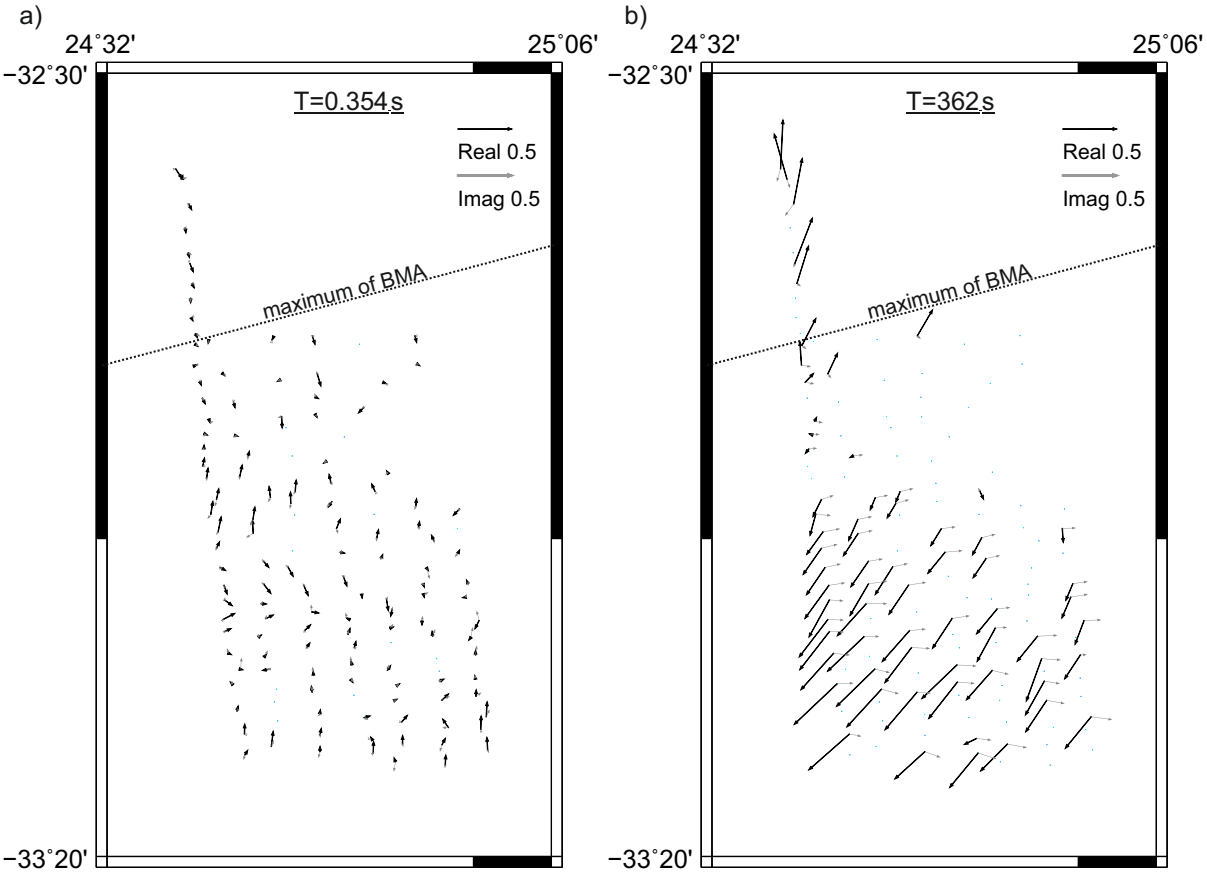

Figure 3.11: Map of induction vectors for the two representative periods a) $T = 0.354\,s$ and b) $T = 362\,s$.

In general, we observe for most MT stations that induction vectors for short periods ($T < 10\,s$) are small and close to zero (Fig. 3.11a). This implies a layered or homogeneous conductivity structure in the upper part of the subsurface. For longer periods, the magnitude of the induction vectors increases and they deviate significantly from zero (Fig. 3.11b).

A prominent reversal can be observed for real induction vectors: North of the surface trace of the BMA maximum they point in north and northeast direction, while south of it they point in southwest direction. The induction vectors in Figure 3.11b indicate a shift of the BMA maximum in southern direction. Especially in the southern part, the imaginary induction vectors are not parallel to the real induction vectors indicating a 3D conductivity structure.

In the framework of this thesis, 34 inter-station transfer functions could be calculated with station 312 as the reference station. The station locations of the corresponding 34 local stations are displayed in Figure 3.6 and exemplary curves are shown in Figure 2.21. Almost all inter-station transfer functions differ significantly from the identity matrix for the shortest periods ($T < 1/512\,s$). Short periods have a low penetration depth and differences in this period range can indicate that the conductivity structure for the upper metres or tens of metres is locally different for the stations. For the long periods ($T > 32\,s$) several smaller deviations from the identity matrix, representing a layered subsurface, can be observed for the real part of M_{xx} depending on the station location. For stations in the south, the real part drops from 1 to 0.7 whereby this effect decreases to the central part. Stations in the central part resemble the identity matrix for this period range supporting the hypothesis that parts of the study area can be explained with a layered electrical conductivity structure. For stations in the northern part, the real part of M_{xx} increases from 1 up to 1.2. Deviations from the identity matrix for the longest periods are a hint of a more complex electrical conductivity structure in greater depth. For the period range between $1/512\,s$ and $32\,s$, most of the inter-station transfer functions resemble the identity matrix.

In summary, all available transfer functions indicate a layered electrical conductivity structure in the upper part. For longer periods as well as in the southern part, the structure becomes more complex. This rough estimate of electrical conductivity structure seems to correspond well with a layered sedimentary basin, which shows a more complex geology in the south due to influence of the Cape Fold Beld orogeny.

So far, the MT data set from 2005 (Weckmann et al., 2007a) could only be interpreted in terms of 2D inversion. 2D inversions of the new data set including a strike direction analysis as well as a detailed geological interpretation has been done by Greve within the framework of her master thesis (in preparation). Now, an additional 3D inversion study is aimed due to the newly acquired stations. Within the scope of this thesis, the resolution capacities of 3D models for a very thin potential shale gas bearing formation within a resistive sedimentary basin was tested. The aims were (i) to verify if 3D inversion can provide additional information compared to 2D inversions or if the result is mainly controlled by the regularisation, (ii) to detect possible problems and (iii) to derive information of the

electrical conductivity of the target horizon, the Whitehill Formation.

Chapter summary:

- The Karoo Basin is a retro-arc foreland basin representing about 100 million years of sedimentation.
- Large amounts of potentially recoverable shale gas are estimated in the black shales of the Whitehill Formation.
- Previous studies showed that the Whitehill Formation is a marker horizon with high conductivities.
- In 2014, a MT study was conducted with the aim to map the Whitehill Formation.
- Transfer function estimation of these data could significantly be improved by the novel criteria.
- The transfer functions imply a layered electrical conductivity structure in the upper part, a high conductive feature in the depth and a more complex structure for longer periods and in the southern part.

4 Inversion

The MT method aims to determine the subsurface's electrical conductivity structure by measuring natural variations of electric and magnetic fields at the Earth's surface. To get a conductivity distribution with depth from the transfer functions, forward modelling and inversion tools have to be applied.

In recent years, 3D inversion techniques have become available and practical due to advancement in computational resources. Some 3D modelling programmes based on various approaches have been developed for both forward modelling and inversion (e.g. Siripunvaraporn et al., 2005a,b; Egbert & Kelbert, 2012) that are freely available for academic use. The focus of this thesis was to investigate whether an areal coverage and subsequent 3D inversion of the data set foster a more realistic mapping of physical properties of the target horizon, a potential shale gas bearing formation. I finally would like to derive provisions in terms of experimental layout, favourable input data for inversion and inversion strategies. 2D modelling is done by J. Greve within the framework of her master's thesis (Greve, in preparation), which finally allows to compare the results of 2D and 3D inversion in terms of feasibility and appropriateness for exploration.

In this thesis, I used the Modular Electromagnetic Inversion system ModEM (Meqbel, 2009; Egbert & Kelbert, 2012; Kelbert et al., 2014) for 3D modelling. The ModEM package contains programmes and routines for both forward modelling and inversion of frequency domain EM data with gradient-based search methods. The 3D MT modelling scheme applies a finite difference approach to solve Maxwell's equations numerically. Furthermore, a non-linear conjugate gradients algorithm is used within the inversion process to solve the minimisation problem.

In the following section, I briefly summarise the 3D inversion of MT data with ModEM. A more detailed description can be found in Meqbel (2009); Egbert & Kelbert (2012); Kelbert et al. (2014).

4.1 Basic description of 3D inversion of MT data with ModEM

In forward modelling, the propagation of EM wave fields for a known discretised subsurface is simulated allowing the calculation of transfer functions at any position in the model. For solving the MT forward problem, a large number of model parameters $\mathbf{m} = (m_1, m_2, \ldots, m_M)^T$ are required, which describe

the Earth's electrical conductivity structure. The so-called model responses, e.g. transfer functions, are represented by the data vector $\mathbf{d} = (d_1, d_2, \ldots, d_N)^T$. The forward problem

$$\mathbf{d} = F(\mathbf{m}) \tag{4.1}$$

with F being an in general non-linear forward operator has always a unique solution.

In the inversion problem, the model parameters $\mathbf{m}$ have to be derived from the data $\mathbf{d}$, which are represented e.g. by the measured data. This is done through an iterative process that modifies the current model until a variety is found that explains the data within pre-defined error bounds. As the number of data parameters N is much smaller than the number of model parameters M, the inversion problem is ill-posed and has no unique solution.

For the solution of the non-linear inversion problem, model parameters have to be found that explain the data, e.g by minimising an objective or penalty function Φ :

$$\Phi(\mathbf{m}, \mathbf{d}) = (\mathbf{d} - F(\mathbf{m}))^T {\mathbf{C}_\mathrm{d}}^{-1} (\mathbf{d} - F(\mathbf{m})) \tag{4.2}$$

with $\mathbf{C}_\mathrm{d}$ being the covariance matrix of data errors.

The term on the right-hand side of equation (4.2) is a measure of data misfit, i.e. the difference between measured data and model responses.

As the inversion problem is ill-posed and underdetermined, normally model regularisation terms $\Omega(\mathbf{m})$ are introduced to stabilise the problem and to limit the number of possible solutions:

$$\Phi(\mathbf{m}, \mathbf{d}) = (\mathbf{d} - F(\mathbf{m}))^T {\mathbf{C}_\mathrm{d}}^{-1} (\mathbf{d} - F(\mathbf{m})) + \lambda \Omega(\mathbf{m}) \tag{4.3}$$

with the regularisation parameter λ. This parameter controls the influence of the regularisation term during the inversion and furthermore determines the trade-off between data fit and model regularisation. In ModEM, λ is determined by an automatic criterion and decreases during the inversion. At the beginning, the inversion is mainly controlled by model regularisation, whereas towards the end the inversion is driven by the data misfit.

In MT, the regularisation term is often formulated in a way that the inversion searchs for (i) a smooth model, where the transition between conductivity structures varies smoothly and (ii) a model that is as close as possible to a prior model. ModEM combines both approaches and the model regularisation

term $\Omega(\mathbf{m})$ is defined as:

$$\Omega(\mathbf{m}) = (\mathbf{m} - \mathbf{m}_0)^T \mathbf{C}_m^{-1} (\mathbf{m} - \mathbf{m}_0) \tag{4.4}$$

with $\mathbf{m}_0$ being a set of prior model parameters and $\mathbf{C}_m$ being the positive definite symmetric model covariance. This model regularisation prevents models that are too rough and/or are far from the prior model. Furthermore, the inversion keeps the assumed prior resistivities in model areas which are poorly constrained. Often the prior model is equal to the starting model for the optimisation. However, ModEM also allows to define them differently.

The model covariance matrix is calculated by a sequence of 1D smoothing and scaling operators (Tietze, 2012):

$$\mathbf{C}_m = \mathbf{c}_x \mathbf{c}_y \mathbf{c}_z \mathbf{c}_z^T \mathbf{c}_y^T \mathbf{c}_x^T. \tag{4.5}$$

Each of these 1D smoothing operators is block-diagonal as in the following shown for the $x-$direction:

$$\mathbf{c}_x = \begin{pmatrix} c_{11}^x & & & \\ & c_{21}^x & & \\ & & \ddots & \\ & & & c_{NyNz}^x \end{pmatrix} \tag{4.6}$$

with Ny and Nz cells in $y-$ and $z-$direction, respectively. The blocks c_{jk}^x are constructed by an autoregression scheme using model covariance parameters, e.g the parameter α_x in $x-$direction:

$$c_{jk}^x = \begin{pmatrix} 1 & & & & \\ \alpha_x & 1 & & & \\ \alpha_x^2 & \alpha_x & 1 & & \\ \vdots & & & \ddots & \\ \alpha_x^{Nx-1} & & & & 1 \end{pmatrix} \tag{4.7}$$

The smoothing parameters α_i lie in the interval $[0, 1]$.

The minimum of the penalty function is determined by a LSQ approach. At each iteration, the non-linear conjugated gradients algorithm within ModEM performs two steps. First, the gradient of $\Phi(\mathbf{m}, \mathbf{d})$ is calculated to obtain the direction in which $\Phi(\mathbf{m}, \mathbf{d})$ decreases most quickly. Secondly, a step size is determined in order to find a local minimum of the penalty function along the direction given by the first step. This procedure is repeated until the inversion reaches one of the predefined stopping criteria, e.g. a minimum data misfit or a maximum number of iterations. Several other parameters and inversion settings can be defined by the user, e.g. discretisation of the model grid,

the starting and prior model, the used data set as well as the corresponding error bounds and the smoothing parameters α_x, α_y and α_z.

The data misfit in ModEM is defined as a global root mean square value:

$$RMS = \sqrt{\frac{1}{N}\sum_{n=1}^{N}\frac{(d_n^{obs} - d_n^{mod})^2}{e_n^2}} \tag{4.8}$$

with $\mathbf{d}^{obs}$ and $\mathbf{d}^{mod}$ being the data vectors of the measured (observed) and the modelled data and e_n^2 being the squared data errors. The RMS in equation (4.8) is defined as a normalised value and hence it is unitless. An RMS value of 1 corresponds to an optimal data fit within the given error bounds. As the RMS value in ModEM combines the data fit for all components, sites and periods, it can only be considered as a rough estimate of inversion result's quality. A visual comparison between measured and modelled data is necessary to assess and to rate different inversion results.

The used 3D modelling routines of ModEM are parallelised (Meqbel, 2009) reducing both memory requirements and run times. All inversions within the framework of this thesis were run on the GFZ computer cluster. The code is parallelised over periods and the two orthogonal source field polarisations. The minimum computation time can be achieved by distributing the inversion problem on $2N_p + 1$ processors with N_p being the number of periods used in the modelling. For my examples, I used only every second period (25 in total) of the data to further decrease the computation cost. The complete data set consists of four component impedance data for 141 stations, vertical transfer functions (VTF) for 135 stations and 34 inter-station transfer functions in a period range of $10^{-4} - 10^{3}\,s$. The masked impedances and VTFs used for the inversions are shown in the right columns in appendix A.4-A.10.

4.2 Deriving a 3D model for the Karoo data set

In a first step, I derived a 3D model for the Karoo data set by finding suitable inversion parameters and combining all available information of the transfer functions. Additional resolution tests were conducted to examine conspicuous electrical conductivity features and to test the resolution depth of the obtained model. In the second step, I examined if the derived 3D model is able to provide information of the physical properties of the target horizon. Parameters of the inversion model that are indicative for physical properties, such as porosity or thermal maturity, are the depth, thickness and electrical conductivity of the target horizon. Furthermore, areas of reduced electrical conductivity

within this horizon or even an absence of high electrical conductivity values are of great interest as they might indicate sweet spots that still contain economically viable and recoverable shale gas. Synthetic case studies were sometimes conducted in addition to test and verify certain aspects.

4.2.1 Model setup

Before inversions were carried out, I evaluated different inversion settings and parameters as well as their influences on the result. The impedance data were used for these tests. Important parameters that were tested in the framework of this thesis are e.g. the grid, the starting model, the data errors and the smoothing parameters. Since a presentation of all inversion tests would go beyond the constraints of this thesis, I will only present an essence of the relevant results and their implications.

Model grid

An initial step in setting up a 3D inversion is the specification of an inversion coordinate system and an appropriate model grid. For all tested grids, I used a coordinate system with $x-$ and $y-$axis pointing towards geographic north and east, respectively. As ModEM utilises a cartesian, right-hand coordinate system, $z-$axis points downwards.

The extensions of the different grids were chosen large enough to cover parts of the Indian and Atlantic Ocean due to the fact that e.g. Meqbel et al. (2014) and Weckmann (2015) showed that the influence of the surrounding ocean should be taken into account. Modelling tests for my study area confirmed that the ocean influences the long period data ($T > 10\,s$) of the different transfer functions. In order to reduce the computation time for such large models, a nested modelling approach (Meqbel et al., 2014) was used for most of the inversions. This nested approach consists of two steps: (i) the induction equations are solved on a coarser grid covering a larger area with realistic ocean/continent geometry; (ii) this solution is used to provide boundary data for a smaller area centred on the actual study area. A number of different model grids were tested with edge lengths of $750\,m$, $1000\,m$ and $1200\,m$ in $x-$ and $y-$direction, whereby the cell size in $y-$direction was always equal or greater than that in $x-$direction due to the different site spacing for the two directions. The selected cell sizes guaranteed that maximal one station lies within a cell. 30 padding cells were added at each horizontal direction with increasing size by a factor of 1.2 in order to avoid boundary effects. The number of cells in $z-$direction varied between $60 - 80$ with increasing cell sizes by a factor of 1.1 or 1.2 starting from a thickness of $10\,m$ or $50\,m$ at the surface. The final inversion results of all models were similar, but the final misfit was slightly different. The finally selected grid consists in total of $140 \times 110 \times 60$ cells in the two horizontal and the vertical direction, respectively. The inner part comprises a uniform mesh

of $80 \times 50 \times 60$ with an edge length of $1000\,m$. The vertical thickness of the first layer was set to $10\,m$. I chose a value of 1.2 for the increasing factor of the padding cells as well as for the vertical cells. The inner grid with additional five padding cells in each horizontal direction was used as nested model.

Starting model

For an optimal inversion result, it is important to find a reasonable starting model. Therefore, different homogeneous half-spaces with $30\,\Omega m$, $100\,\Omega m$, $300\,\Omega m$ and $500\,\Omega m$ were tested. For all of them, I included the bathymetry of the ocean with $0.3\,\Omega m$ (fixed value) as a-priori information. The best inversion result and data fit was observed for the $500\,\Omega m$ starting model.

Data errors

The data errors are an important parameter in the inversion (e.g. Tietze, 2012). For the VTFs the estimated variances were combined with an error floor of 2 % resulting in larger errors for larger induction vectors. A constant error of 0.02 was applied to the inter-station transfer functions. Different approaches and error bounds were tested for the impedance components, e.g. a combination of estimated variances and error bounds as well as fixed errors. This was important, because unfavourable selected errors can lead to completely wrong inversion results, e.g. poorly determined models, which are unable to fit the data and to resolve the target formation. Normally, the estimated variances are large for the long period data due to the small amount of available events. To better account for the long period data by reducing the data errors for this period range, I decided to use fixed errors. These errors could be computed using percentages of e.g. $\sqrt{|Z_{xy} \cdot Z_{yx}|}$, $|Z_{ij}|$ or a combination of both. Percentage values of $3-5\,\%$ for the off-diagonal and $20-100\,\%$ for the main-diagonal elements were tested. Best data fit could be obtained for 5 % and 50 % of $|Z_{ij}|$ for the off- and main-diagonal components, respectively.

Smoothing parameters

Further important parameters are the smoothing values α_x, α_y and α_z. Various different smoothing parameter combinations were examined with $0.1 \leq \alpha \leq 0.4$. The smoothing in $y-$direction was always equal or larger than that in $x-$direction due to the larger site spacing in $y-$direction and the smoothing in $z-$direction was always smaller or equal to the horizontal smoothing. A smoothing value of 0.2 in each direction provided the best compromise between minimal overall misfit and a reasonably smooth model.

4.2.2 Inversion of individual transfer functions

First, the impedance tensor data and the VTFs were separately inverted to get an impression of their influence on the overall conductivity image of the study area.

Impedance tensor inversion

The electrical conductivity model derived from the impedance data (Fig. 4.1) has large similarities with the previous 2D inversion result from Weckmann et al. (2007a).

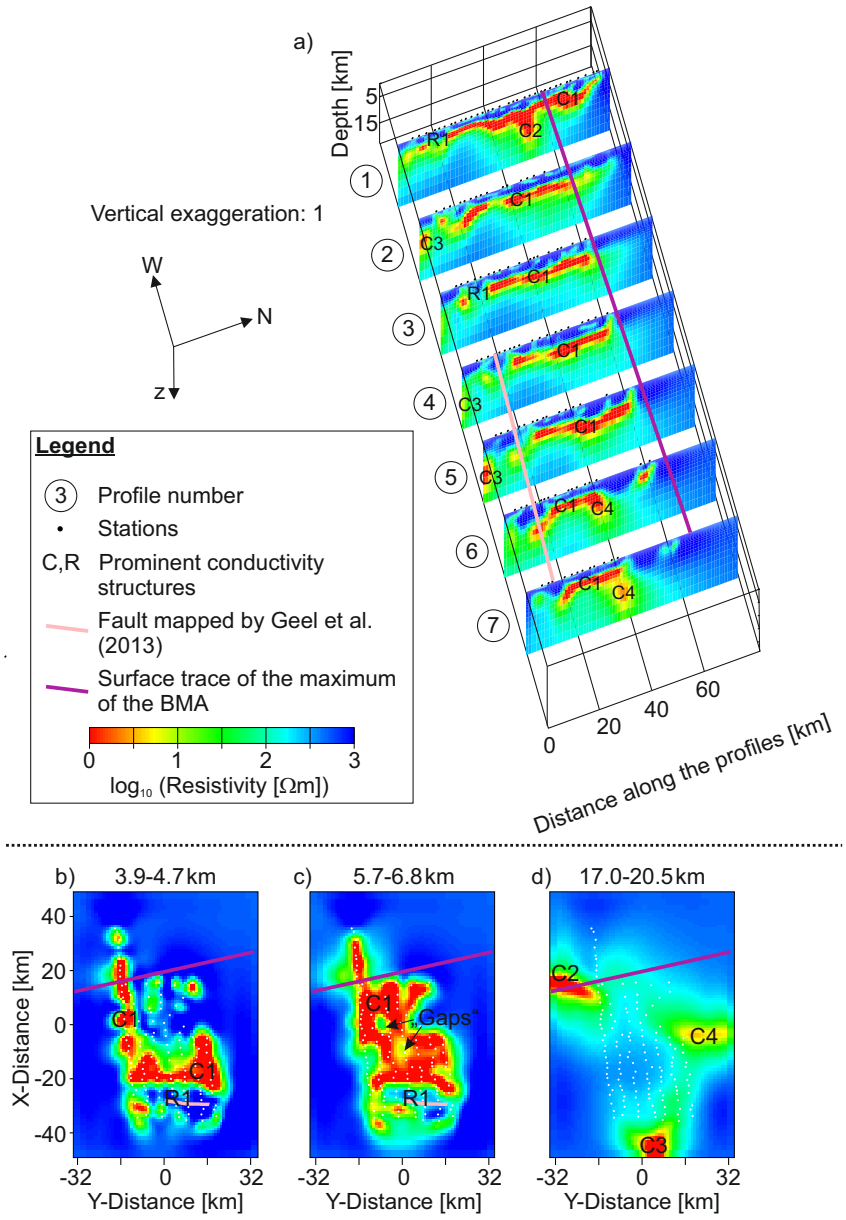

Figure 4.1: Results of the impedance inversion: a) Electrical resistivity cross sections along the seven profiles. The target horizon is associated with the conductor C1. b)-c) Different depth slices showing the spatial extent of C1. d) Depth slice indicating additional electrically conductive structures in greater depth. The overall conductivity structure and the labelled features are discussed in more detail in the text.

In general, the resolution outside the station coverage is relatively low resulting in only small changes from the starting model in these areas. The upper $2\,km$ of the subsurface (except from the first few metres) are mainly dominated by high resistivities $> 500\,\Omega m$. A horizontal layer of high electrical conductivities is the most prominent feature in the model labelled with C1 (Figs. 4.1a-c). This conductive layer seems to correlate with the target horizon, the Whitehill Formation, and has a thickness of several hundreds of metres in the model. Over large areas, resistivities below $1\,\Omega m$ are observed for C1, reaching even resistivity values up to $0.1\,\Omega m$ and below in some parts. Although this electrically conductive layer is more or less continuous for the entire area, there exist smaller gaps with increased electrical resistivity values (Fig. 4.1c). Furthermore, in the southern part this layer seems to be interrupted by an area of higher resistivities (R1).

A second conductive feature (C2) is visible in the northern part of profile 1 directly below the high conductive layer C1. This vertical conductor extends further to the west as shown in Figure 4.1d and deepens in western direction. Conductor C2 matches well with the vertical conductor seen by Weckmann et al. (2007a) directly beneath the surface trace of the BMA maximum. In profile $2-6$, a conductive feature (C3) is indicated in larger depths in the southernmost part of the profiles. Another deep conductor (C4) is resolved by profile 6 and 7 in the eastern part of the study area. The conductors C3 and C4 as well as the feature R1 will be discussed later in this chapter.

The observed result corresponds well with the previous study and the qualitative data interpretation in the previous chapter. Moreover, the model fits the data to an RMS of 1.16 (start value: 49.11) after 143 iterations, whereby especially the off-diagonal components could be well fitted (Fig. 4.2).

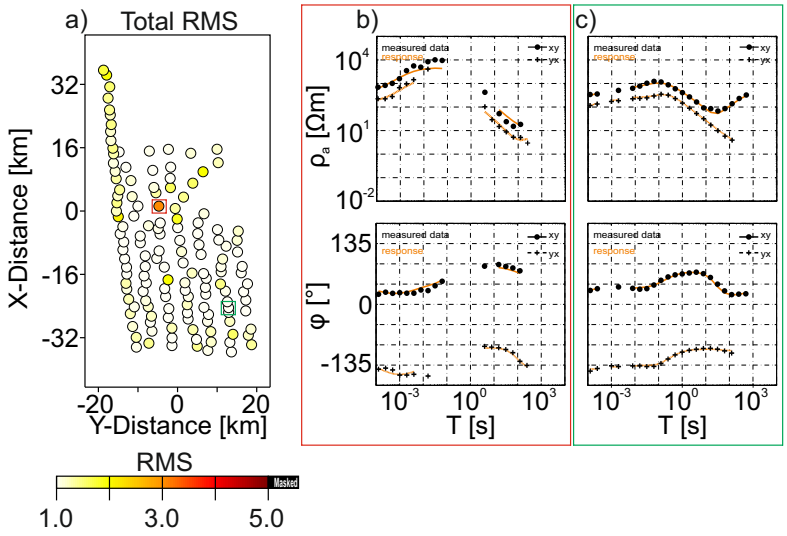

Figure 4.2: Data fit for impedance tensor inversion: a) Overall data fit for all stations. b)-c) Comparison of measured and predicted off-diagonal tensor components for the two stations highlighted in a).

Inversion of vertical transfer functions

The inversion result of the VTFs (Fig. 4.3) shows only small changes from the starting model and thus contains much less structure than the model derived from the impedance data.

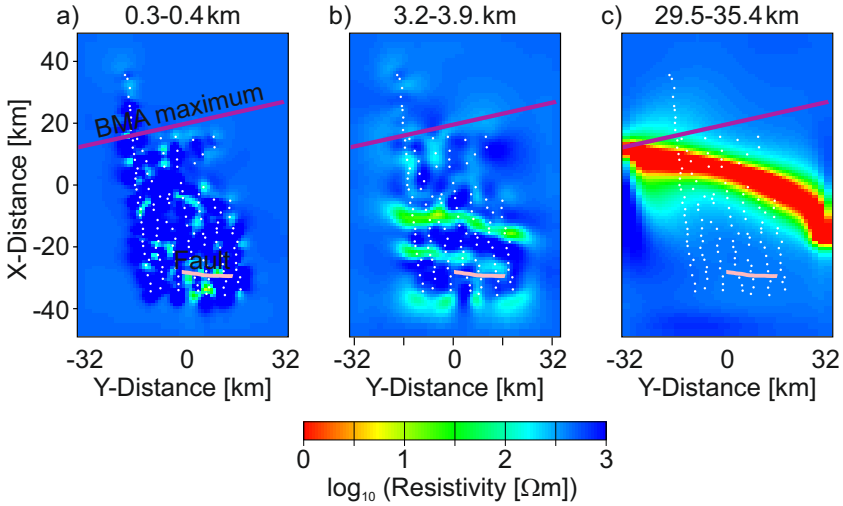

Figure 4.3: Different depth slices from the inversion of the VTFs overlain by a fault mapped by Geel et al. (2013) and the surface trace of the maximum of the BMA. The most prominent feature is a band of high electrical conductivities (c).

Several east-west trending zones of increased electrical conductivity exist between 2 km and 6 km in the southern and central part of the study area (Fig. 4.3b). In greater depth ($17 - 50\,km$), the most important feature is a band of high conductivities running from east to west (Fig. 4.3c). Additional tests with different starting models confirmed the position and existence of the electrically conductive band, whereby the depth of this conductor could be decreased with decreased resistivity of the starting model.

In the western part of the model, the conductive band coincides from its location with the conductor C2 and in the eastern part with the conductor C4 of the impedance inversion. C2 was interpreted as a vertical conductor beneath the maximum of the BMA (Weckmann et al., 2007a,b). However, from overview maps the surface trace of the maximum of the BMA would run from west to northeast (e.g. Fig. 4.3c). Therefore, an explanation for the east-west trending band of high electrical conductivities is not obvious and can only be given after further modelling studies.

The inversion stopped after 45 iterations with a final misfit value of 1.05 (starting misfit: 3.54).

4.2.3 Synthetic case study to test different transfer function combinations

After inverting individual transfer functions, I want to combine all available information of the different transfer functions in one model and increase the resolution capacity by performing a joint inversion. Before I performed a joint inversion with real data, I conducted a synthetic case study to examine the influence of the different transfer functions in joint inversions.

For the synthetic study, I used the real station layout and a simplified model derived from the two individual inversions. As I applied the same forward solver for (i) calculating the forward responses and (ii) the subsequent inversion, I changed at least the model grids and added 5 % and 2 % Gaussian noise to the impedances and the VTFs as well as inter-station transfer functions. The ability to recover different structures was tested by performing several 3D joint inversions using different combinations of the three different transfer functions.

The synthetic model

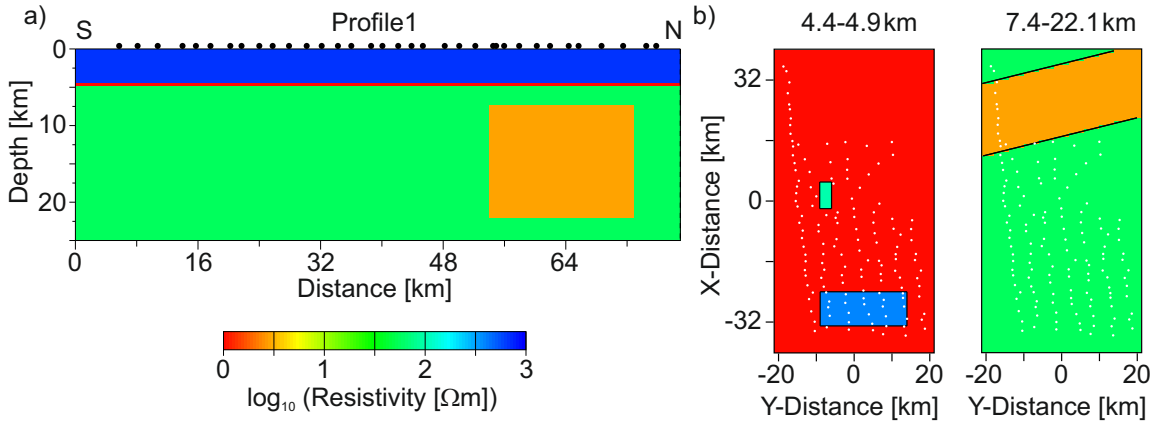

Figure 4.4: a) Cross section along profile 1 and b) two representative depth slices of the synthetic model 1 described in more detail in the text.

The synthetic model 1 (Fig. 4.4) has a background consisting of three layers following the key structures of the impedance inversion model: (i) a resistive layer of 700 Ωm in the upper 4 km representing the typical resistive Karoo sediments, (ii) a highly contiguous conductive layer of 1 Ωm representing the conductive Whitehill Formation and (iii) a conductive layer of 60 Ωm up to 100 km representing average resistivities observed directly below the conductive layer in the inversion model of the impedance data. A half-space of 20 Ωm was added at the bottom of the model to ensure that EM

fields for the considered period range do not penetrate away from the utilised grid (Campanyà et al., 2016). Several 3D structures were integrated into this 1D background model: (i) a resistive body of 400 Ωm that interrupts the conductive layer in the southern part of the study area, (ii) a smaller gap of 100 Ωm in the conductive layer in the northwest where the station coverage is insufficient and where the conductive layer is interrupted in the impedance inversion of the field data and (iii) a vertical conductor of 3 Ωm between $7.4 - 22.1\,km$ representing the conductor C2 below the maximum of the BMA. In addition, a second model similar to the first one was created without the smaller gap of 100 Ωm. The forward responses for all possible components of the three different transfer functions were calculated for 139 stations and 25 periods in a range of $10^{-4} - 10^{3}\,s$ on a large grid consisting of in total (including padding cells) $140 \times 100 \times 80$ cells in north-south, east-west and vertical direction. The inner part comprises $100 \times 60 \times 80$ cells with edge lengths of $1000\,m$. Twenty padding cells in each horizontal direction increase with a factor of 1.3. The thickness of the top layer in $z-$direction was set to $50\,m$ with an increasing factor of 1.1.

Inversions were performed for all individual transfer functions as well as for all possible different transfer function combinations. To allow an unbiased comparison of differences due to the particular data type used for inversion, the same inversion parameters were used for all inversions. The grid used for the inversions was coarser than the previously described forward grid. It consists of $113 \times 80 \times 60$ cells with an inner part of $83 \times 50 \times 60$ cells. The edge lengths of the core cells are $1200\,m$ and the padding cells increase with a factor of 1.3. The first layer is $50\,m$ thick with an increasing factor of 1.15. The dimension of this grid corresponds well to the grid used for the real data, even if they are not identical.

Similar to the real inversions, the optimal inversion parameters were found after different tests. The parameters and the final settings are summarised in Table 4.1.

First, a starting model was chosen after running impedance tensor inversions with different half-space models. Ten different starting models were tested covering a large spectrum of possible starting values. The inversion result using a 150 Ωm half-space obtained the most similar conductivity values to the given synthetic model, although for all inversions similar RMS values were achieved. For these tests, a smoothing value of 0.2 was used for all three directions following the final smoothing values for the inversions of the real data.

After finding an appropriate starting model, I tested different smoothing parameters, because the regularisation is important and has an influence on the final inversion result. As different data types can react differently to the same parameter settings, tests for the smoothing parameters were executed

for impedances and a joint inversion of all three available transfer functions. The final parameters were set to 0.3 for the two horizontal directions and to 0.1 for the $z-$direction.

Parameter	Tested setting	Final setting
Starting model (homogeneous half-space)	$20 - 700\,\Omega m$	$150\,\Omega m$
Smoothing parameters $(\alpha_x, \alpha_y, \alpha_z)$	$0.1 - 0.6$	$\alpha_x = \alpha_y = 0.3\ \alpha_z = 0.1$
Reference station for inter-station transfer functions	• above a 3D structure (station 306, 407, 409) • next to a 3D structure (station 223, 320, 404) • on top of a 1D environment (station 312, 515, 708)	on top of a 1D environment (station 312)

Table 4.1: Summary of the tested parameters for the synthetic case study 1. More details are described in the text.

An additional test was performed to find an adequate reference station for the inter-station transfer functions. Campanyà et al. (2016) showed that the choice of the reference station influences the convergence of the inversion process as well as the resolution capacity. The tested stations were located directly above or next to the 3D structures as well as on top of a normal 1D environment. The obtained inversion models showed that locating the reference station on top of a 1D environment better recovers the original synthetic model. Using a reference station close or directly above a 3D structure either led to an unstable inversion or resulted in a poorly resolved model that was not able to explain the data. During the field experiment, station 312 was selected as reference station due to its high data quality. The synthetic case study confirms that this was also an excellent choice with regard to use it for the inter-station transfer functions as it was located in a normal 1D environment. Therefore, I decided to use this station also as reference station for the synthetic case study.

Results

Several individual and joint inversions were conducted using impedance tensor, VTFs and inter-station transfer functions. I tested if and how VTFs and inter-station transfer functions contribute in a joint inversion to increase the resolution capacity. A good inversion result should (i) recover the true subsurface structure as good as possible, (ii) should be able to fit the data and (iii) introduce as little

as possible artefacts. Here, I refer only to four of these different inversions (Figs. 4.5 & 4.6).

All four inversion models could fit the data to a similar level. The layered background structure consisting of the resistive top layer and the electrically conductive layer was recovered in all models (Fig. 4.5). Furthermore, the two resistive structures within the electrically conductive layer could be fully recovered.

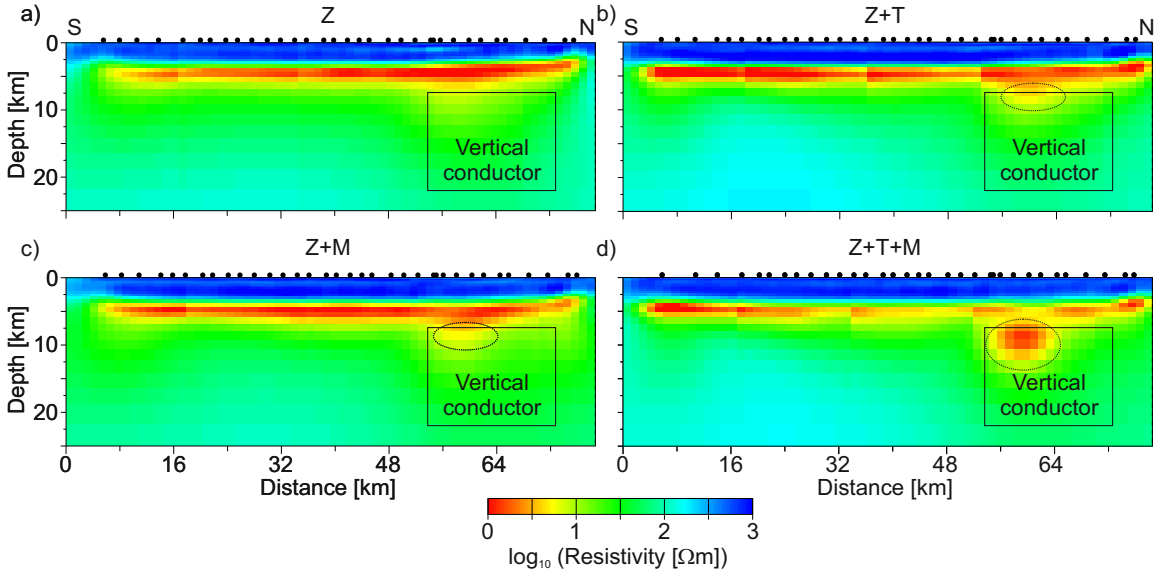

Figure 4.5: Cross sections along profile 1 for four inversions using different transfer function combinations for synthetic model 1. Although all models could recover the layered background structure, only the combination of all three transfer functions was able to recover the electrical conductor (indicated by the black box) beneath the BMA maximum. The dashed ellipses highlight increased electrical conductivities associated with this conductor.

However, the conducted study suggests that it is difficult to recover the vertical conductor beneath the BMA maximum indicating that the resolution below the electrically conductive layer is low. Inversion of only impedance data results in a model where this vertical conductor is missing (Fig. 4.5a). Joint inversions of impdance tensor data and VTFs (Fig. 4.5b) or inter-station transfer functions (Fig. 4.5c) indicate this vertical conductor by slightly increased electrical conductivities directly below the conductive layer. Nonetheless, this vertical conductor can only be reasonable recovered by using the VTFs and inter-station transfer functions in addition (Fig. 4.5d). For this joint inversion, the vertical conductor is explicitly resolved and it starts to separate from the overlying conductive layer. In addition, all inversion models do not introduce any major artefacts.

As the target horizon of this study is the electrically conductive layer, I also examined its resolution. All shown inversions (Figs. 4.5 & 4.6) recover this layer. Furthermore, this layer is in all models thicker than in the true model. This shows that even a thin layer will always appear thicker in the

inversion model e.g. due to the vertical smoothing. The greatest thickness is observed if only the impedance data are inverted.

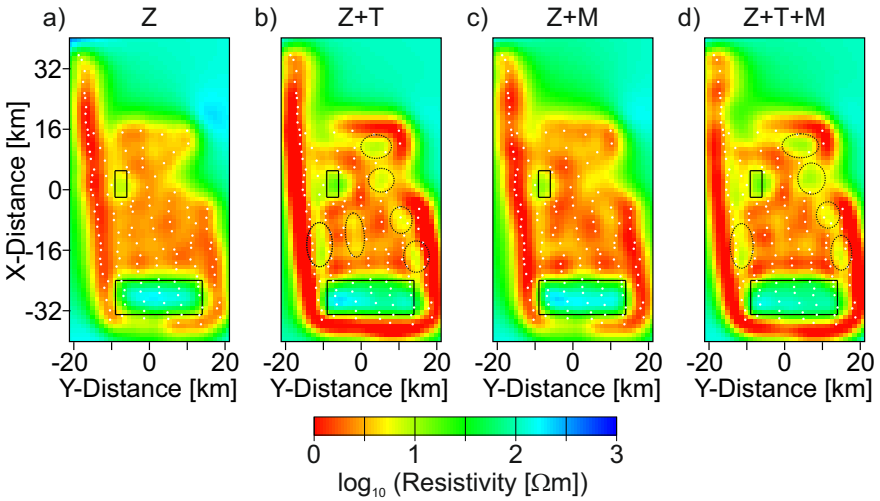

Figure 4.6: Depth slices of $4.4 - 5.1\,km$ for the same four inversions as in Figure 4.5. Combinations of impedance data with VTFs and/or inter-station transfer functions better recover the true conductivity of the electrical conductive layer. The dashed ellipses highlight artificial gaps.

Performing a joint inversion of impedance data with VTFs and/or inter-station transfer functions better recovers the true conductivity of the electrical conductive layer (Fig. 4.6) and results in a thinner electrically conductive layer. Although the inversion results using either VTFs or inter-station transfer functions are similar along profile 1 (Figs. 4.5b & c), the spatial resolution of the electrially conductive layer differs for these two models (Figs. 4.6b & c). Impedance and inter-station transfer functions (Fig. 4.6c) recover this layer almost contiguous with only smaller areas of slightly decreased conductivity, whereas the joint inversion of impedance and VTFs (Fig. 4.6b) results in a less contiguous layer that has several smaller disruptions (gaps). These disruptions are in particular observed between the stations of adjacent profiles and suggest, that inter-station transfer functions can be used as an alternative to VTFs in cases where measurement of vertical magnetic fields was not possible. Artificial gaps are also observed for the second synthetic model with a more contiguous conductive layer, if VTFs are used in a joint inversion. The artificial gaps occur also for the joint inversion of all three transfer function types (Fig. 4.6d), which is on the other hand the only inversion that was able to recover all structures of the true model. As areas of increased resistivity within the electrically conductive layer could be interpreted as potential shale gas bearing spots in the exploration context, it is important to know that smaller gaps could also be an artefact of the used transfer function type and/or the chosen station layout.

4.2.4 Joint inversions and the preferred model for the Karoo data set

Based on the insights of the synthetic case study, inversion of the real data was conducted accordingly. In contrast to the synthetic data, the error of real data is unknown and not all transfer functions exist for every station. This holds particularly for the inter-station transfer functions that are available for only 34 stations.

Inversion strategy

First, I performed joint inversions by using impedance data with either only VTFs or only the inter-station transfer functions (results not shown). A similar subsurface electrical conductivity structure as for the impedance inversion was recovered for both models. Larger differences were observed for the vertical conductor C2 and for the conductive layer C1.

The joint inversion using the VTFs results in a model where the conductor C2 goes down to greater depths and where the conductive layer C1 is less contiguous than in the inversion of only impedance data. Both aspects correspond well with the observations of the synthetic case study. The data fit of the impedance data as well as for the VTFs was slightly worse in comparison to the data fit of the individual inversions. This applies in particular to long period VTF data. To improve data fit and to increase the weight of the long period data during inversion, I decided to use only the long period ($T > 64\,s$) VTF data. This is justified, as for most of the stations only data within this period range differ significantly from zero. A joint inversion with this reduced data set for the VTFs improves the data fit of both transfer functions and results in a more contiguous electrically conductive layer C1.

The joint inversion of impedance and inter-station transfer functions recovers the conductor C1 as a contiguous layer, but has problems to resolve the vertical conductor C2. This is very likely due to the small data set of inter-station transfer functions that are all focused in the central part. Similar to the joint inversion of impedance and VTFs, the data fit is worse than for the individual impedance inversion. Furthermore, especially the short periods of the inter-station transfer functions could not be fitted. As the focus of this study was to resolve the potential shale gas bearing horizon which is expressed in periods greater than $1/100\,s$, I removed the short periods ($T < 1/1000\,s$) for the inter-station transfer functions.

In another study building on these results, I combined impedances (entire period range), VTFs ($T > 64\,s$) and inter-station transfer functions ($T > 1/1000\,s$). Although the overall misfit (from 49.62 to 1.95) was reasonable, a closer look at the data fit of the different transfer functions and stations

revealed, that in particular the long period data were not adequately fitted (Fig. 4.7a).

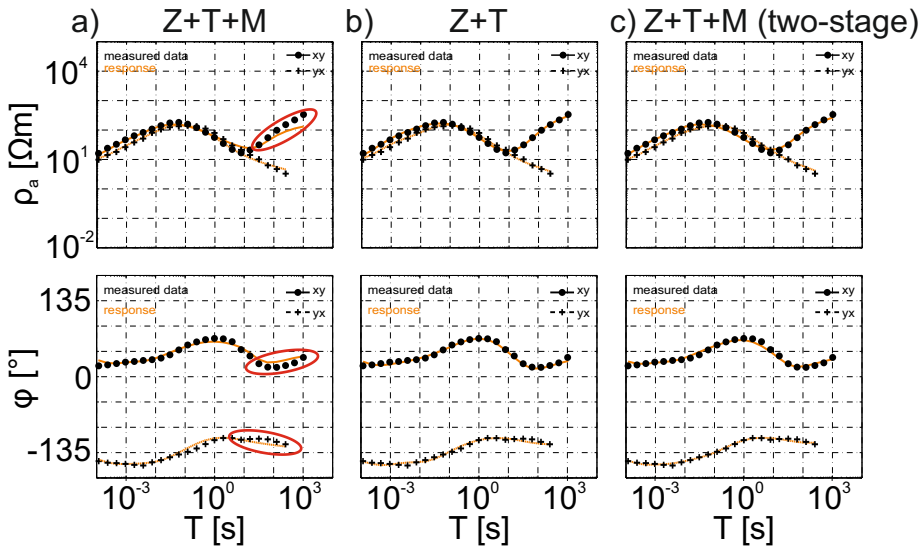

Figure 4.7: Data fit of the off-diagonal impedance tensor components for an exemplary station from the south for different joint inversions. Using the two-stage approach (described in the text) improves the data fit for joint inversions using all three transfer function types for the components and periods highlighted by the red ellipses.

I aimed at increasing the data fit in particular for the long period data, as this period range ($T > {}^{1}/_{100}\,s$) very likely contains information about the target horizon and underlying layers. For this reason, I used two steps to combine all available transfer functions: (i) I inverted only impedance and VTFs ($T > 64\,s$) starting from a $500\,\Omega m$ background model. Thereby, I was able to obtain a sufficient data fit (Fig. 4.7b) for these two transfer functions (RMS: from 41.16 to 1.5). (ii) I used the inversion model of the first step as starting model for a joint inversion using now the inter-station transfer functions ($T > {}^{1}/_{1000}\,s$) in addition. With this two-stage approach, I received a better data fit for all transfer functions for the long period data in comparison to the joint inversion using all three transfer functions in one go (Fig. 4.7). The overall misfit of this final model is 1.88 (start rms: 2.00). The overall misfit for the impedance data as well as a map of real induction vectors comparing measured and predicted data for one representative period are shown in Figure 4.8. Using only the inter-station transfer functions in the second step was not possible as the inversion became unstable. As a consequence, the two-stage inversion result is quite similar to the joint inversion result of impedance and VTFs. As I had only 34 inter-station transfer functions, their impact on the inversion was significantly lower than from the other two transfer functions. Nevertheless, due to inter-station transfer functions, the electrically conductive layer became more contiguous for areas that are located between local and reference station. Therefore, I prefer the inversion result of the two-stage inversion approach for the Karoo data set. This model is the basis for further examinations

to test the resolution capacity of the target horizon.

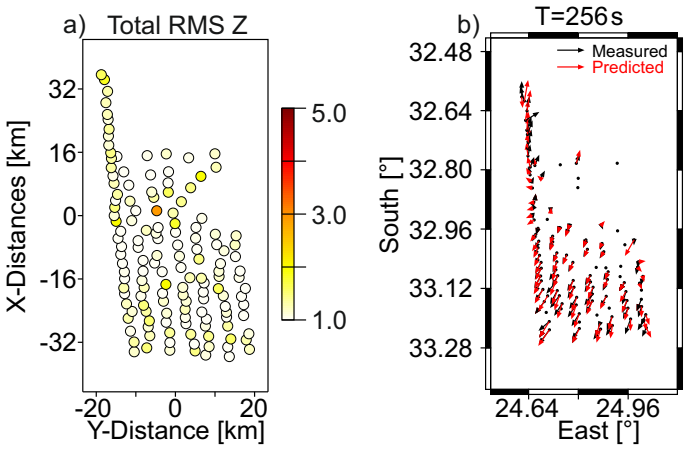

Figure 4.8: Exemplary data fit of the preferred model. a) Overall mistfit for the impedance tensor data. b) Overview map showing real induction vectors comparing measured and predicted data for one representative period. For both transfer functions the misfit is sufficient.

Preferred 3D model

The result of the preferred joint inversion (Fig. 4.9) resembles in general the inversion result of the impedance data (Fig. 4.1): The upper two kilometres of the subsurface are dominated by higher resistivities followed by a layer of high electrical conductivities (C1). The high resistivities in the upper $2\,km$ correspond well with the resistivities given by van Zijl (2006) and Weckmann et al. (2007a). They are caused by the Karoo sediments. Smaller conductive features within these layers were not evaluated in the framework of this thesis, but they can be indicators of shallow aquifers.

Below these resistive sediments a horizontal conductive layer with resistivities $\leq 1\,\Omega m$ can be associated with the carbon rich Whitehill (see van Zijl, 2006; Branch et al., 2007; Weckmann et al., 2007a). The observed high conductivities have been interpreted in terms of black shales with a very high thermal maturity close to the meta-anthracite field (Branch et al., 2007). This thermal maturity suggests that the organic matter is solely consumed in the rock matrix and available gaseous filling is either consumed or left the matrix. Therefore, no producable shale gas can be expected in this area. In comparison to the impedance inversion (Fig. 4.1), the layer C1 is less contiguous in the preferred model (Fig. 4.9) and shows more and greater gaps of increased resistivity (Fig. 4.9c). These gaps are of particular interest as they could indicate spots of lower thermal maturity that still contain some shale gas. However, such conclusions can only be drawn, if these gaps are not caused by inversion settings, station coverage or data availability.

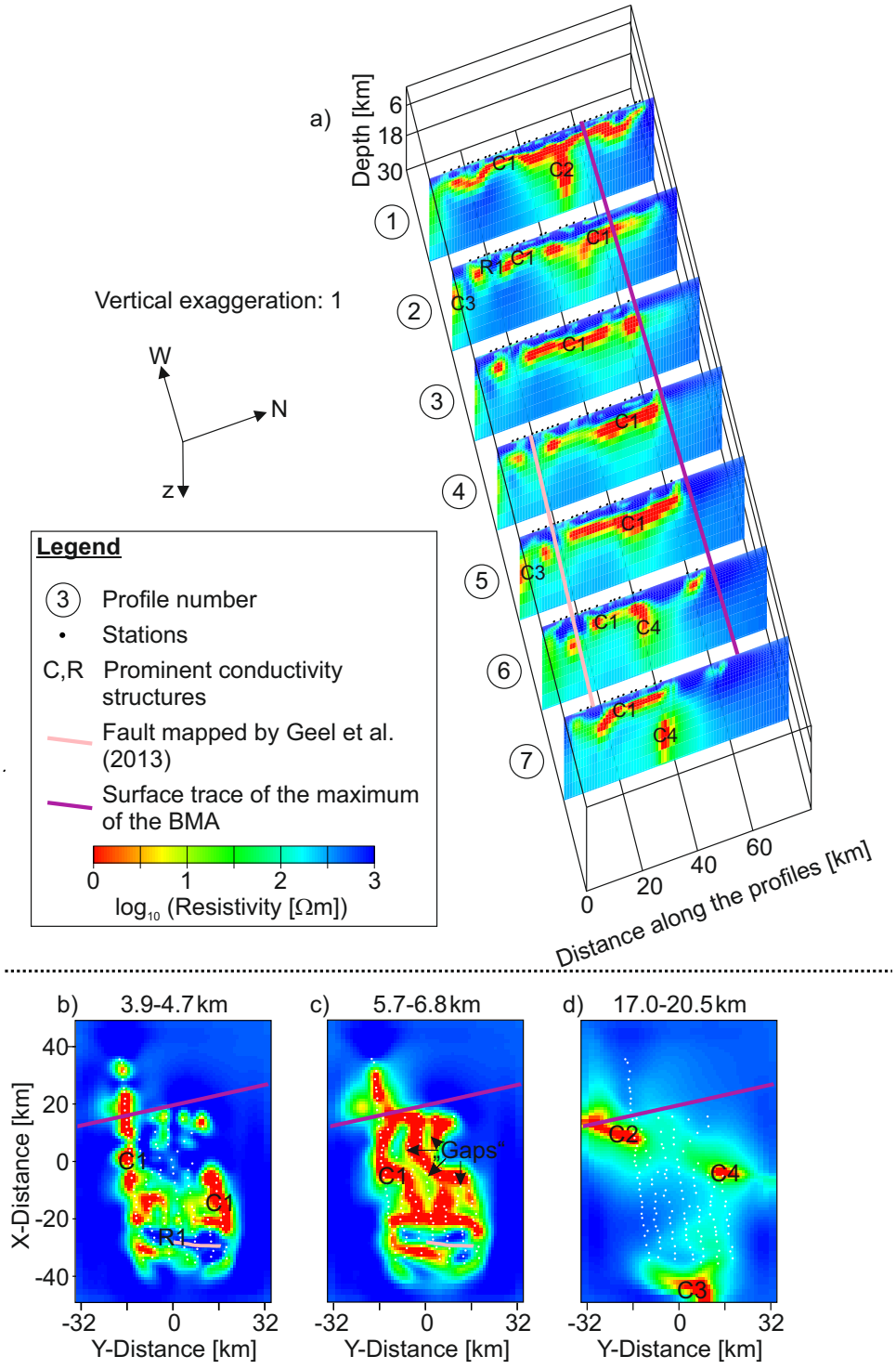

Figure 4.9: Results of the preferred model: a) Electrical resistivity cross sections along the seven profiles. The target horizon is associated with the conductor C1. b)-c) Different depth slices showing the spatial extent of C1. d) Depth slice indicating additional electrically conductive structures in greater depth. The conductor C1 seems less contiguous and many smaller gaps are visible.

Moreover, the target horizon has a thickness of several hundred metres and is located between $3-8\,km$.

The obtained thickness is far beyond the thickness of the relevant layers retrieved by nearby well logs.

To study the properties of the target horizon resolved by 3D inversion, I first tested if this layer could be thinned without worsening the data fit. For derivation of physical properties of the target horizon from the preferred model, it is important that the corresponding parameters such as e.g. depth, resistivity and thickness are well determined. Therefore, additional tests were performed.

The conductor C2 is clearly imaged (e.g. profile 1 in Fig. 4.9a) and extends to greater depths than for the impedance tensor inversion (Fig. 4.1). The conductive feature C4 in the east shows similarities with the conductor C4 in the impedance tensor inversion but extends, as C2, to greater depths. From the inversion model of the VTFs, that connects these two conductors, can be assumed that these conductors are caused by the same structure. The conductor C3 in the south is still resolved, but is located more to the south and appears therefore smaller in profiles $2-5$. The resolution of these three additional conductive features was tested in the next section to decide whether they are inversion artefacts or are required by the inverted data.

The feature R1, that interrupts the conductive layer in the south, has higher resistivities as in the impedance inversion. Furthermore, the electrically conductive layer seems to continue south of this structure (Fig. 4.9c). The resistive feature R1 is located in the area where the Whitehill Formation is bent or somehow off-set and thus is found closer to the surface. In none of the profiles the conductive layer could be traced from the depth to the surface. The observed higher resistivities might be caused by weathering or a lower thermal maturity of the shallower parts of this formation. Furthermore, the southern edge of R1 coincides with a fault mapped by Geel et al. (2013) and Geel (2014). As the feature R1 continues to the west this indicates that the fault extends to the west.

4.2.5 Resolution study of the preferred model

Before I analysed different parameters of the target horizon in more detail, I tested the depth resolution of the shown models and examined whether the conductors C2-C4 are required by the data.

Resolution depth of the data

The depth resolution of the impedance data and of the preferred model was tested using a "squeeze test" as described in Meqbel et al. (2014). For this test, the resistivity below a specific depth was set to $100\,\Omega m$ and was fixed. The inversion was started using the modified model and the impact of the constraint on the data fit and the model structure was examined. Three different depths were tested covering a range from $24-50\,km$.

For the impedance data, the RMS value was the same for all three tests and no significant changes

were observed in the models. This suggests that the impedance data are not sensitive to structures below 24 km. For the joint inversion, this kind of modification in a depth of 24 km results in a larger start RMS value and the model changes in particular in the eastern part around the conductor C4. Further tests showed that the combination of all data is sensitive to conductivity structures down to 30 km.

Resolution of conductors C2-C4

The resolution of the three conductors C2-C4 was tested separately with several constrained inversions. For this purpose, the conductors were removed from the preferred model by replacing the high conductivities with average values of the surrounding area and the model was restarted.

These tests suggest that conductor C2 is needed to explain the data. This result was to be expected as Weckmann et al. (2007a) already imaged a vertical conductor in mid-crust depth below the maximum of the BMA. The preferred model recovers this conductor (C2) and its extent to the western part. Due to insufficient station coverage in large parts along the surface trace of the BMA's maximum, it could not be successfully recovered in the eastern part of the study area. Another reason why the vertical conductor is not resolved below the BMA maximum in the eastern part might be that the shear zone that causes the vertical conductor does not always run along the BMA maximum. However, if the shear zone was located more in the south, it would be resolved by the data.

The conductor C3 in the south neither reappeared in these tests nor I observed any additional conductive structures outside this area; but the data fit was slightly worse for impedance as well as VTFs. This suggests that the conductor C3 was caused by a structure outside the study area and the dimensions of the nested grid. This could be the influence of some sedimentary basins further in the south, e.g. Algoa and Gamtoos Basin. Similar high conductivity basins were also observed farther west in the Cape Fold Belt by Weckmann et al. (2012). There, the Oudtshoorn Basin was extremely conductive hosting a reservoir of possibly hot and saline fluids. Unfortunately, electrical conductivity values and thicknesses of the Algoa and Gamtoos Basin are poorly determined (Jade Greve, pers. comm.).

The overall misfit increases if conductor C4 is removed although the data fit worsens only at some stations for long period data. The conductor reappeared in these tests indicating that this conductor is caused by a real structure. The conductor C4 was connected with the conductor C2 by a band of high electrical conductivities in the inversion of the VTFs (Fig. 4.3) suggesting that both conductors are caused by the same structure. To verify this hypothesis, additional synthetic studies were conducted.

Synthetic case studies to analysis the conductor C4

For this test, I used the same synthetic model as described in section 4.2.3 (Fig. 4.4). This model already contains a vertical conductor below the surface trace of the BMA's maximum. A 3 Ωm body is embedded in the 60 Ωm background along the surface trace in a depth of $7.4 - 22.1\,km$.

Depth slices showing the maximal extension of this conductor for different transfer function combinations (Fig. 4.10) show that inversions using only the impedance data (Fig. 4.10a) cannot recover this conductor. This has been shown before in section 4.2.3 (e.g. Fig. 4.4a). The reason for this is the lower depth resolution of the impedance data due to the overlying electrically conductive layer. The best resolution of this conductor could be achieved by using only the VTFs (Fig. 4.10b) as VTFs are very sensitive to lateral conductivity contrasts. For this inversion, the conductor follows the given shape, it is recovered for almost the complete depth range (not shown) and it is resolved from the western edge up to profile 4. This result emphasises the ability of the VTFs to detect and map lateral conductivity contrasts even if they are located outside the station coverage.

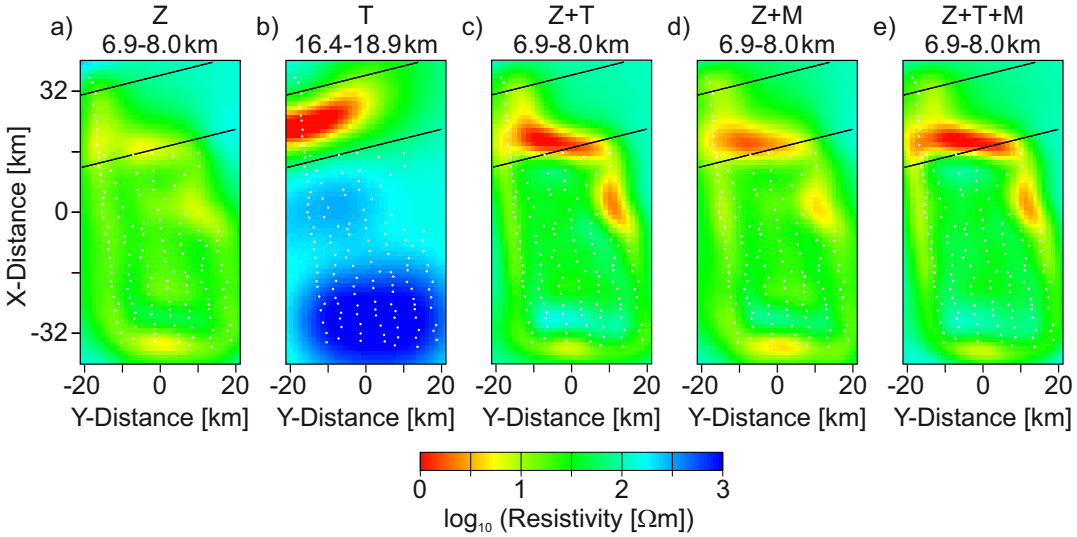

Figure 4.10: Different depth slices to show the resolution of the conductor below the surface trace of the maximum of the BMA for synthetic model 1 using different transfer function combinations. The black lines indicate the extent of this conductor in the true model. The vertical conductor could be recovered for most of the models, but its location and shape differs between the models.

Joint inversions using VTFs and/or inter-station transfer functions in addition to the impedance data result in very similar models (Figs. 4.10c-e). For all of these three models, a vertical conductor is resolved along the northern edge of the station coverage and a smaller conductor appears in the east where the station coverage has a gap. This conductor in the east is located close to the position of C4 for the real data. However, its goes only down to 8.5 km. In contrast, the conductive feature in the north extends down to 12.9 km that comes closer to the real value. This conductor (Fig. 4.10c-e)

is rotated in southeastern direction in comparison to its original shape in the synthetic model (Fig. 4.4b). However, a band of high conductivities running from west to east, as observed in the inversion model of the VTFs (Fig. 4.3c), cannot be simulated with this synthetic model. That might be because the data basis for this synthetic study and the real inversions differs: For the synthetic study I used a "perfect" data set that contains all three different transfer functions at each station and for each period. In contrast, the real data set is much smaller due to some technical problems and insufficient data quality for some transfer functions and stations. For this reason, I created a second synthetic case study that uses a more realistic model and data set.

The model of the second synthetic study was still kept simple, but it contains more realistic structures than the first synthetic model. Furthermore, the same model setup and data structure as for the real inversions were used to simulate the real inversion conditions as good as possible. That means that transfer function components were only calculated if they exist for the real data resulting in a reduced data set in comparison to the data set used for the first synthetic model.

For the forward calculation, an inner grid consisting of $110 \times 70 \times 77$ cells was used with edge lengths of $750\,m$ and 35 padding cells in each horizontal direction with an increasing factor of 1.2. The first layer is $10\,m$ thick and the cell thickness increases with a factor of 1.15.

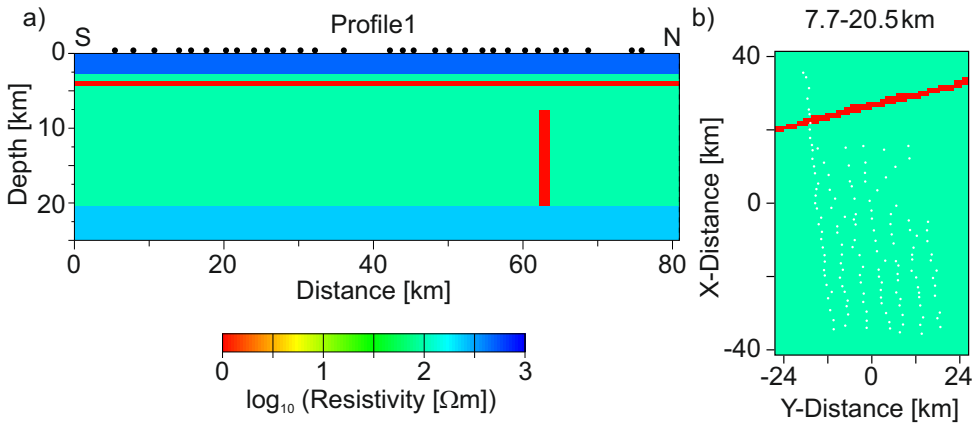

Figure 4.11: a) Cross section along profile 1 and b) representative depth slice of the synthetic model 2.

The background model (Fig. 4.11) consists of three layers: (i) $100\,\Omega m$ for the first $20\,m$, (ii) $500\,\Omega m$ between $0.02 - 2.8\,km$ and (iii) $100\,\Omega m$ between $2.8 - 20\,km$. A half-space of $250\,\Omega m$ was added for depths greater than $20\,km$. Additional 3D structures were integrated in the background model to create a more realistic model based on e.g. previous inversion results (Figs. 4.1 & 4.3): (i) the conductive ocean with a resistivity of $0.3\,\Omega m$, (ii) a conductive body of $1\,\Omega m$ along the entire Karoo

Basin at a depth of $3.8 - 4.3\,km$ representing the conductive Whitehill Formation, (iii) a gap in this conductive body with a resistivity of $700\,\Omega m$ in the southern part of the study area and (iv) a vertical conductor of $1\,\Omega m$ along the surface trace of the maximum of the BMA in a depth of $7.7 - 20.5\,km$. The assumed conductivity of $1\,\Omega m$ for this vertical conductor comes closer to the values observed in the preferred model than the $3\,\Omega m$ body of the first synthetic model. In return, the vertical conductor has a smaller and more realistic width. Again, the conductive layer representing the target horizon and the vertical conductor are separated by more than $3\,km$.

Inversions of different transfer functions were conducted using a nested modelling approach. As mentioned above, the same inversion grid as for the inversions with real data was used. In addition, the same inversion parameters were applied, starting the inversion from a $500\,\Omega m$ half-space, including the ocean bathymetry as a-priori information and setting the smoothing parameters to 0.2 for all three directions.

The inversion results using only the VTFs as well as performing a joint inversion of VTFs and inter-station transfer functions (without impedance data) are shown in Figure 4.12. Both models image a band of high conductivities in east-west direction, whereby the recent result (Fig. 4.12b) looks very similar to the real data inversion result (Fig. 4.3c). Differences between the two models in Figure 4.12 are probably caused by different sensitivities of the different transfer functions, but are not studied in more detail in the framework of this thesis.

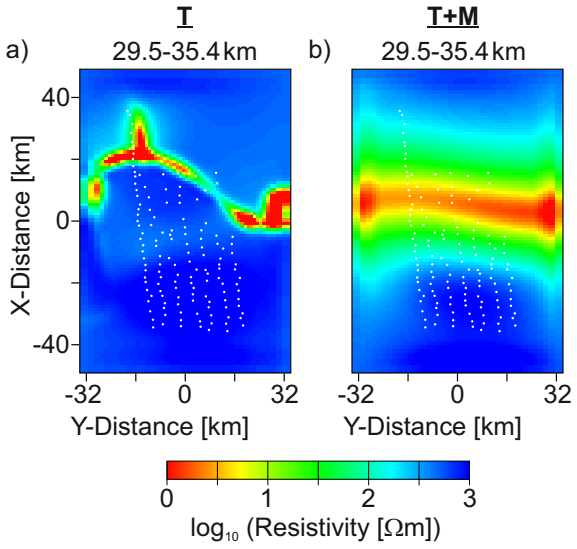

Figure 4.12: Depth slices from the inversion of a) VTFs and b) VTFs and inter-station transfer functions for synthetic model 2 showing a roughly east-west running band of high conductivities. This synethetic study suggests that the band of high electrical conductivities that is deviated towards the south is likely caused by a projection of a conductive feature in the north.

These synthetic case studies suggest that the observed conductive band in the inversion result of the VTFs (Fig. 4.3c) as well as the conductor C4 in the preferred model (Fig. 4.9) are caused by a vertical conductor below the surface trace of the maximum of the BMA in the northern part that is projected into the study area. This projection was possible because the station coverage in the eastern part is less dense than in other parts of the study area. Furthermore, the data quality of the VTFs was low in the northern part. To achieve better results in the future, a denser station coverage (without larger, irregular gaps) and additional data (i.e. VTFs and inter-station transfer functions) are necessary.

4.3 Analysis of different parameters of the target horizon

After deriving a robust 3D model for the Karoo data set that utilises all available information, I analysed different parameters of the target horizon. From an exploration perspective, at least two aspects are important: (i) the conductivity of the target horizon as a proxy for e.g. thermal maturity and porosity and (ii) the variability in conductivity to identify potential so-called sweet spots that still contain gas. A third interesting aspect is the thickness of the layer. Well logs from nearby drillholes reveal that the Whitehill Formation is $\sim 30\,m$ thick in the study area. However, with a normal grid such thin cell sizes in a depth of $4\,km$ are neither feasible nor would a diffusive method such as MT be able to unambiguously resolve the exact properties of this layer. Nevertheless, I tried to thin out the target horizon in the model and examined the influence on the conductivity.

4.3.1 Conductivity of the target horizon

The high electrical conductivity values between $2.2 - 8\,km$ are related to the target horizon. To get an impression of the absolute electrical conductivity values in this depth range, I investigated the resistivity values of each cell without considering the different cell dimensions. After Van Zijl (2006) the resistivity of the Whitehill Formation covers a range between $0.01 - 10\,\Omega m$. This values were derived from many deep electrical sounding curves that their able to resolve the Whitehill Formation as a very conductive layer. In addition, some of the sounding curves were calibrated at e.g. boreholes. Furthermore, Weckmann et al. (2007a) observed similar resistivity values up to $0.3\,\Omega m$ in their constrained inversion. Similar resistivity values are observed for the conductive layer in the preferred model. Most of the cells related to this conductive layer have a resistivity value between

$0.01 - 10\,\Omega m$. Extremely conductive values below $0.01\,\Omega m$ were only observed for very few cells (less than 1 %). Thus the observed conductivity values are in good agreement with previous studies. The high electrical conductivity values suggest that large parts of the target horizon have a high thermal maturity and that no producable shale gas can be expected in these areas. Therefore, areas of increased resistivity within the target horizon are of great interest. They could indicate sweet spots that still contain some shale gas.

4.3.2 Potential areas of lower electrical conductivity

The conductive layer associated with the target horizon has many spots of decreased electrical conductivity in the preferred model (Fig. 4.13c). In contrast, inversion of only impedance data (Fig. 4.13a) or joint inversion of impedance and inter-station transfer functions (Fig. 4.13b) image this layer as more contiguous.

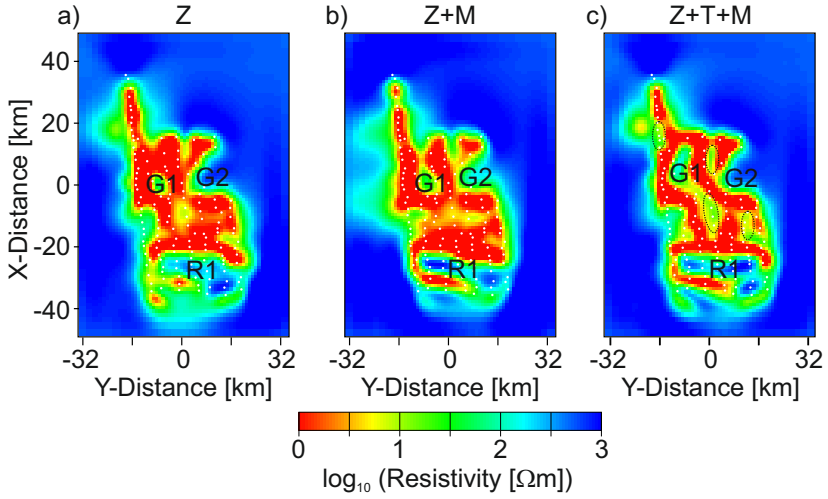

Figure 4.13: Depth slices of $5.7 - 6.8\,km$ for different real data inversions showing the variability in the conductive layer. For the preferred model (c) the target horizon is less contiguous as for the a) impedance and b) joint inversion of impedance and inter-station transfer functions. The labels G1, G2 and R1 highlight spots of decreased conductivity that occur in all inversions. The dashed ellipses in c) show additional gaps in the preferred model.

This raises the question of whether the observed gaps are real or just inversion artefacts. Some of these gaps (highlighted in Fig. 4.13) are consistent for all inversions. The largest gap is the resistor R1 in the south. The structure R1 is well covered by stations and is necessary to explain the data. It was already discussed in the previous section. Its location coincides with the area where the target horizon comes to surface and with a knwon fault.

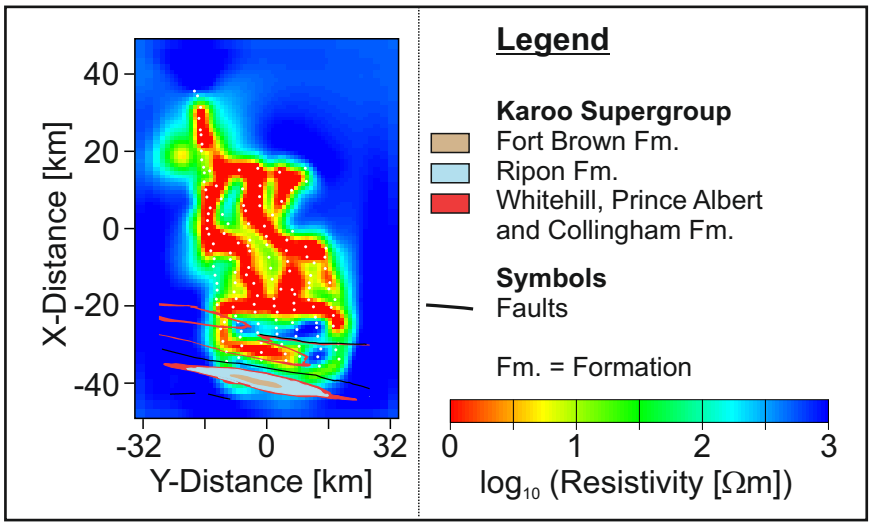

Figure 4.14: Depth slice of $5.7 - 6.8\,km$ of the preferred model overlain by known faults and i.a. the target formation. The red lines represent areas where the target formation outcrops.

In addition, two smaller gaps exist in all models (G1 and G2 in Fig. 4.13). They are located between stations of adjacent profiles indicating that these structures could be caused by unfavourable station layout. This hypothesis is supported by the fact that the synthetic study (Fig. 4.6) produced an artificial gap at a similar position as G2. In additional resolution tests using impedance and inter-station transfer functions, the gaps were closed by using conductivities of $1\,\Omega m$ and the inversions were restarted. Connecting the electrical conductivity layer has no influence on the data fit and the conductive layer remained contiguous in these areas (results not shown). Thus, both data types are not sensitive to the gaps G1 and G2.

Additional spots that are observed in the preferred model, are very likely structures unsupported by the data, in this case the VTFs, as suggested by the synthetic case study in 4.2.3 (e.g. Fig. 4.6). It therefore can be concluded that the observed "sweet spots" in the preferred model are solely caused by our field setup, except for the features R1 in the south. Consequently, the target horizon is dominated by high electrical conductivities in the study area suggesting that it is unsuitable for shale gas exploration. I recommend for future studies with a similar focus to make greater use of the inter-station transfer functions as they have a similar resolution capacity for lateral contrasts as VTFs, but introduce less questionable structures in the target horizon. Therefore, inter-station transfer functions are a good alternative to VTFs. Furthermore, a more regular station distribution can reduce the amount and size of these structures.

4.3.3 Thickness of the target horizon

In the inversion models (Figs. 4.1 & 4.9), the inversion usually smeared out the conductive layer over several cells in vertical direction. This is expected since the regularisation also works in z-direction. Furthermore, the MT method is not sensitive to the layer's thickness, but only to the conductance as the product of a layer's conductivity and its thickness. In addition, MT is also not sensitive to the lower boundary of a good conductor. To demonstrate that the modelled "Whitehill Formation" can be thinner, I therefore conducted several constrained inversions.

In a first attempt, I inverted only the impedance data with the previous described inversion grid and settings. To allow sharp conductivity contrasts that enable the inversion to thin the conductive layer, I added a penalty cut in the starting model that interrupts smoothing between cells above and below this cut. The depth of this penalty cut varied between $2 - 8\,km$ for different inversions to test several possible depth ranges that are associated with the conductive layer. For all these inversions, the misfit was much higher than for the normal impedance inversion without a penalty cut. Especially the long period data, that contain information about our target horizon, could not be fitted.

In a second attempt, I used the inversion result of the preferred model as a basis. To constrain the thickness of the conductive layer, I added so-called "tear zones" that include parts of the conductive layer C1 that have a resistivity below a certain value. Normally, the inversion algorithm penalises sharp conductivity contrasts. In contrast, the inversion algorithm allows sharp conductivity contrasts between neighbouring grid cells along tear zone. Moreover, the resistivities of the affected grid cells can still be changed during the inversion. Cells outside of the tear zones and within the relevant depth range ($>2\,km$) were set to a background resistivity of $50\,\Omega m$ and the inversions were started with the modified starting model. In contrast to the first attempt, now the starting model already includes a thin conductive layer and the tear zones allow a more realistic shape of the conductive layer than the horizontal penalty cuts. In this way, I was able to thin the Whitehill in a similar way as Weckmann et al. (2007a) and to achieve a good data fit. For the exemplary model presented here (Fig. 4.15), I used the top of the target horizon determined by cells with resistivity values below $1\,\Omega m$. The corresponding cells kept their true resistivities and were sandwiched by tear zones. The joint inversion of the preferred model was restarted using this modified starting model (Fig. 4.15c). The result shows that the conductive layer C1 could be thinner as it has been resolved in the preferred model (Fig. 4.15b). Only in the northern and southern part, the data are not in agreement with either such a thin conductive layer or require a continuation of higher electrical conductivities, because the inversion introduces additional zones of high conductivities. Thinning of the conductive layer, results

also in an increase of the electrical conductivities within this layer. Now, 50 % of the cells related to the high electrically conductive layer have resistivity values below $1 \Omega m$.

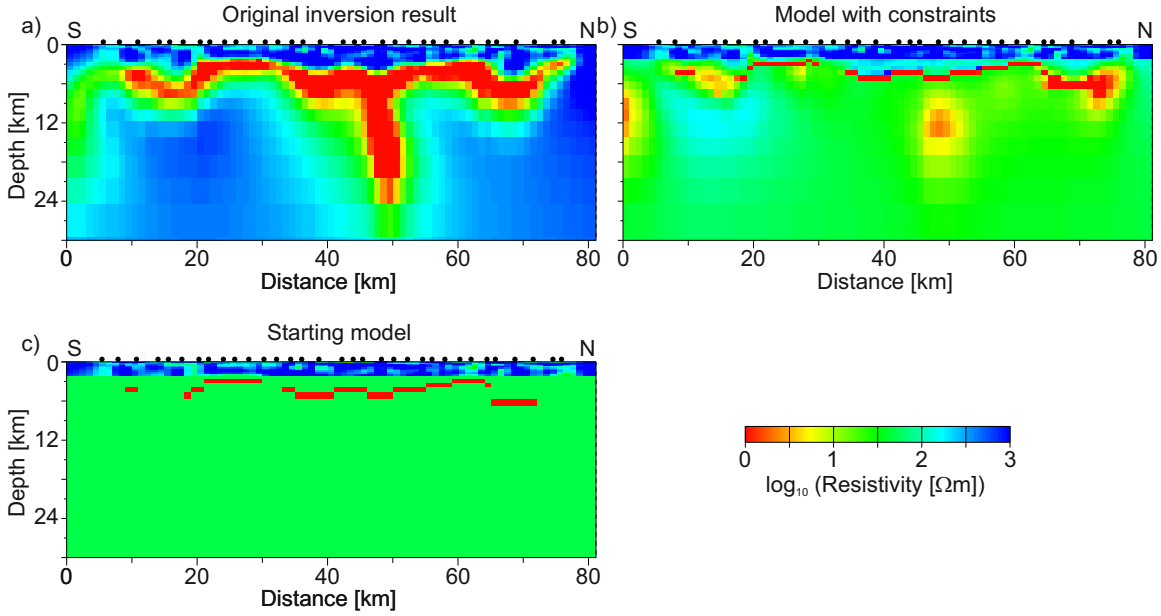

Figure 4.15: Inversion results of a) the preferred model and b) a constrained inversion along profile 1 demonstrating that the conductive layer can be thinner. c) Starting model for the constrained inversion derived from the preferred model. Tear zones enclose regions with resistivity values below $1\,\Omega m$ and all other cells in the relevant depth range ($> 2\,km$) are set to a background resistivity of $50\,\Omega m$.

Furthermore, the result indicates that although this layer seems to be contiguous, it is not perfect horizontal. The depth of the top of this layer varies between $2 - 6\,km$. Moreover, this model suggests that the vertical conductor C2 below the maximum of the BMA is separated from the target horizon. Weckmann et al. (2007a) already discussed the separation of both conductors, also required by geological constraints, on the basis of 2D inversion. With several additional 3D resolution tests, I was able to confirm the required separation.

4.4 Chapter Summary

Experimental layout:

- The reference station for the inter-station transfer functions should be located above a 1D environment to obtain robust and reliable models.
- Uneven station spacing (as observed in the north of my study area) likely causes a subhorizontal conductor at mid-crustal depth co-located with the BMA maximum to be projected towards areas of insufficient station coverage.
- A more regular station distribution could reduce the amount and size of gaps in the conductive target horizon that often occur between more distant stations.

Favourable input data:

- Synthetic case studies suggest that the resolution capacity is increased for joint inversions combining impedance, VTFs and inter-station transfer functions.
- They also indicate that VTFs and inter-station transfer functions have a similar resolution capacity for lateral conductivity contrasts.
- A contiguous electrically conductive layer can be best resolved by a joint inversion of impedance + inter-station transfer functions rather than by a joint inversion of impedance + VTFs that causes artificial gaps.

Inversion strategy:

- As the amount of available inter-station transfer functions was limited, all available data were used to increase the resolution capacity.
- Some data were removed for the VTFs ($T < 64\,s$) and inter-station transfer functions ($T < {}^{1}/_{1000}\,s$) in order to improve the data fit and to focus on period ranges that contain information of the target horizon.
- A two-stage approach was used to combine the different data types as a direct joint inversion of all data types results in an insufficient data fit.

5 Summary

The magnetotelluric (MT) method is useful to image the subsurface's electrical conductivity structure contained in the MT transfer functions and to derive information related to e.g. resources and tectonic processes. Unfortunately, nowadays man-made electromagnetic (EM) noise sources often disturb the measured data significantly and finally hamper the estimation of MT transfer functions, e.g. the impedance tensor. In the past, many advanced techniques, e.g. robust statistics, remote reference processing or noise removal in time and frequency domain, have been suggested to tackle this problem. However, often these approaches are still not sufficient to obtain meaningful results. Depending on the approach, their application can be tedious, time consuming or subject to the user's experience.

Two novel automatic approaches were tested and implemented within the framework of this thesis: (i) a pure Mahalanobis distance (MD) criterion to confine the data to an ideally noise-free subset that is subsequently used in the stacking algorithm and (ii) a selection criterion based on the magnetic polarisation direction (MPD) as a physical add-on for the statistical MD criterion. There exist other robust processings that also uses the concept of the MD (e.g. Egbert, 1997), but here the MD is used as a pre-stack tool to reject single data points belonging to e.g. a separate noise cluster. While the first criterion is statistically limited to a maximum amount of noise of 50 %, the physical add-on can deal with much higher noise contaminations that even exceeds the 50 %.

The distribution of all data was studied to determine the influence of noise on this distribution. In general, there exist two possibilities how noise contribute to the data distribution: (i) single data points scatter around the true MT distribution or (ii) noise forms a separate data cluster that can either be spatially separated from or merged with the desired MT cluster. In addition, outliers can exist that are normally characterised by extreme values. The robust statistics within the standard Emerald processing can remove extreme outliers as well as some of the scattering noise (especially in the tail of the distribution), but fail if the noise contamination is too high and in particular if noise forms a separate cluster.

The MD criterion was developed to confine the data to a subset that contains no or at least significantly less noise, so that the subsequent robust stacking algorithm is able to estimate meaningful transfer functions. The statistical confinement criterion is based on the idea that extreme outliers and noise have a larger distance to the data distribution than data points caused by the MT signal. As a distance measure, I used the MD that is commonly applied in multivariate statistics to detect outliers. In contrast to the normal Euclidean distance, the MD is an advanced distance measure that considers the shape of the distribution and thus can better describe data sets with different variances in each

spatial direction or which exhibit an internal correlation. This is important, because the transfer functions are estimated from auto- and cross-spectra that are derived from electric and magnetic fields which have different magnitude ranges and units. Furthermore, the complex numbered MT transfer function components do not have to have the same variances for their real and imaginary parts and it cannot be excluded that the MT data, i.e. Fourier coefficients, auto- and cross-spectra or transfer functions, exhibit an internal correlation.

In the framework of this thesis, the real and imaginary parts of the transfer functions are used as input data for the MD criterion. Data points with a large distance are removed from the data set prior to the robust stacking by comparing the distance values with a user defined threshold. This affects events in the tail of the distribution, which would be also removed by the standard robust stacking algorithm, but also noise that e.g. forms a separate cluster. For the calculation of a reliable MD value for each single event, the location and covariance matrix have to be computed in a robust manner. Different location and covariance estimators were tested and finally I implemented a deterministic minimum covariance determinant algorithm. This iterative algorithm uses various initial data subsets determined by different statistical estimators to find an ideally noise-free subset for the location and covariance calculation. A "physical" estimator was added to this set based on inductive processes in MT to increase the amount of independent start values. This estimator makes use of the result of an adjacent period to find an optimal subset for the current period; an approach that is not commonly used in most MT data processings. Events that have passed the MD criterion are subsequently used in the robust stacking procedure within the EMERALD processing scheme.

The MD criterion was tested for various MT data sets from diverse regions and with different noise contaminations. A comparison of standard EMERALD processing results with processing results using the MD criterion reveals that for stations with less than 50 % noise contamination, data quality of different transfer functions could be improved over the entire period range, even in the so-called dead band. Many MT stations fulfil the prerequisite that the natural MT signal is not outweighed by EM noise and moreover EM noise often forms a completely independent cluster of transfer functions. The MD criterion is able to remove such clusters and it can reduce scatter around the desired cluster of MT transfer functions leading to a smaller but more focused subset for the subsequent stacking. Therefore, the MD criterion is a useful measure to improve the transfer function estimation in an automated manner, in particular when no other methods such as remote reference can be applied. Nonetheless, the MD criterion is a purely statistical measure and is limited to cases where the majority of all events is well-behaved. Therefore, it can result in totally misleading transfer functions or does not show any improvements for stations affected by more than 50 % with noise or in cases where

the transfer functions of EM noise overlap the desired MT transfer functions to a large extent. In these cases, some noise has to be manually removed e.g. by physically based data selection criteria or other a-priori information to ensure that the majority of all events is well-behaved. For the future, a graphical user interface and a more interactive processing could help the user to select for chosen periods the "correct" cluster even if it consists of the minority of data points.

The MPD criterion was implemented as a physical add-on for the statistical MD criterion to deal with data of periods whose noise contamination exceeds 50 %. It removes data points that are caused by a magnetic polarised signal and thus often reduces the amount of noise to under 50 %, so that the MD criterion can be successfully applied thereafter. The natural magnetic field is generated by a variety of different sources. Therefore, the incidence directions of the generated magnetic fields vary and a preferred polarisation direction is not expected for the desired MT signal. Originally, the magnetic polarisation direction of the magnetic wave field is one parameter of the interactive selection algorithm from Weckmann et al. (2005). This selection tool enables the selection of outliers based on visual inspection of several physical and statistical parameters. Although such interactive tools are highly recommended if data are strongly affected by noise, their application is often tedious and time consuming. Furthermore, the result depends on the user's experience. In contrast, the MPD criterion presented in this thesis works in an automatic manner and is independent of any user input.

Within the MPD criterion, the polarisation directions of all single events are calculated and evaluated together. Kernel routine of this criterion is an iterative algorithm that finds polarisation directions with an exceptionally large amount of events. Therefore, the actual number of events organised in classes is compared with an expected value of a uniform distribution. All internal thresholds of the MPD criterion are selected in a conservative manner to ensure that events are only removed for stations which are influenced by a strongly polarised magnetic source. Slightly difficult are periods that have a small number of events, because it can occur that too many events are removed.

The MPD criterion was tested for many stations with and without interferences of polarised magnetic noise sources. For stations suffering from one or more distinct polarisation directions, the MPD criterion is highly effective for all period ranges and can remove noise even if its amount exceeds the majority of all events. Nevertheless, the MPD also remove data points caused by the desired MT signal that coincides with a disturbed time segment and that has the same incidence direction as the magnetic polarised noise source. In general, it is difficult to prevent the rejection of "good" events that strongly coincides in their incidence direction and temporal occurrence with magnetic polarised signal caused by a noise source. The effect of "good" data points that coincides with noise signal becomes stronger by the fact that the incidence direction is only evaluated in the interval of

$[-90°, 90°]$ instead of the complete interval from $[-180°, 180°]$. Thus some data points of the true MT signal are erroneously projected to disturbed polarisation directions and time segments. The reduced interval was caused by the application of the simple atan and is consistent with the MPD criterion in the interactive selection tool from Weckmann et al. (2005). Although some "good" data points are removed by the MDP criterion, the amount of accepted events was always still sufficient for the subsequent processing steps. In the future, the amount of wrongly removed data can be further reduced by using the complete range of the MPD.

In the second part of this thesis, the novel criteria were used to evaluate a MT data set measured in the Eastern Karoo Basin in South Africa. As part of the presented work, this MT field experiment was jointly conducted by the German Research Centre of Geosciences and the African Earth Observatory Network. This MT study is part of an extensive research programme to collect information prior to a potential shale gas exploitation. The study area was chosen due to the proximity of two existing shallow boreholes and previous seismic and MT studies. In total, five-component broadband MT data were collected at more than 100 stations.

Three-dimensional (3D) inversions were conducted in combination with complementary synthetic studies to (i) derive an image of the electrical conductivity structure of this region with special focus of the potential shale gas bearing Whitehill Formation and to (ii) analyse whether the derived models allow realistic conclusions of physical properties of the target horizon. The preferred model was obtained by a joint inversion of the three different transfer functions. The derived electrical conductivity structure corresponds well with previous studies: Resistive sediments in the upper kilometres are followed by a horizontal conductive layer associated with the Whitehill Formation. Although the Whitehill Formation outcrops in the southern part of the study area, high conductivities can only be observed below a depth of $2\,km$. In the southern area, where the Whitehill comes to surface, the conductive layer is interrupted and high resistivities are observed. The observed higher resistivities of the shallower part of this formation might be caused by weathering or a lower thermal maturity. Deeper vertical conductors in the northwest and east are related to the Beattie Magnetic Anomaly. In the southern part of the study area, the electrical resistivity structure becomes more complex e.g. due to the proximity of sedimentary basins or the Cape Fold Belt.

The conducted inversion and modelling study showed that 3D inversion is capable of resolving the thin potential shale gas bearing target horizon and to map its spatial extent within the study area. Over large areas the resistivities of this horizon are below $1\,\Omega m$ suggesting a very high thermal maturity of the corresponding black shales. This leads to the assumption that most of the gas has already been expunged. Smaller areas within the depth range of the target horizon having higher resistivities are

very likely caused by the uneven station distribution as well as the used data type and are not indicative for producable gas. Furthermore, several provisions in terms of experimental layout, favourable input data and inversion strategies were derived from this study.

In the future, the finally obtained 3D image of the subsurface's conductivity has to be compared with the 2D results. Moreover, the data set and the derived models can be used to study the shallow aquifer system in this fragile environment or to map blind faults.

Appendix

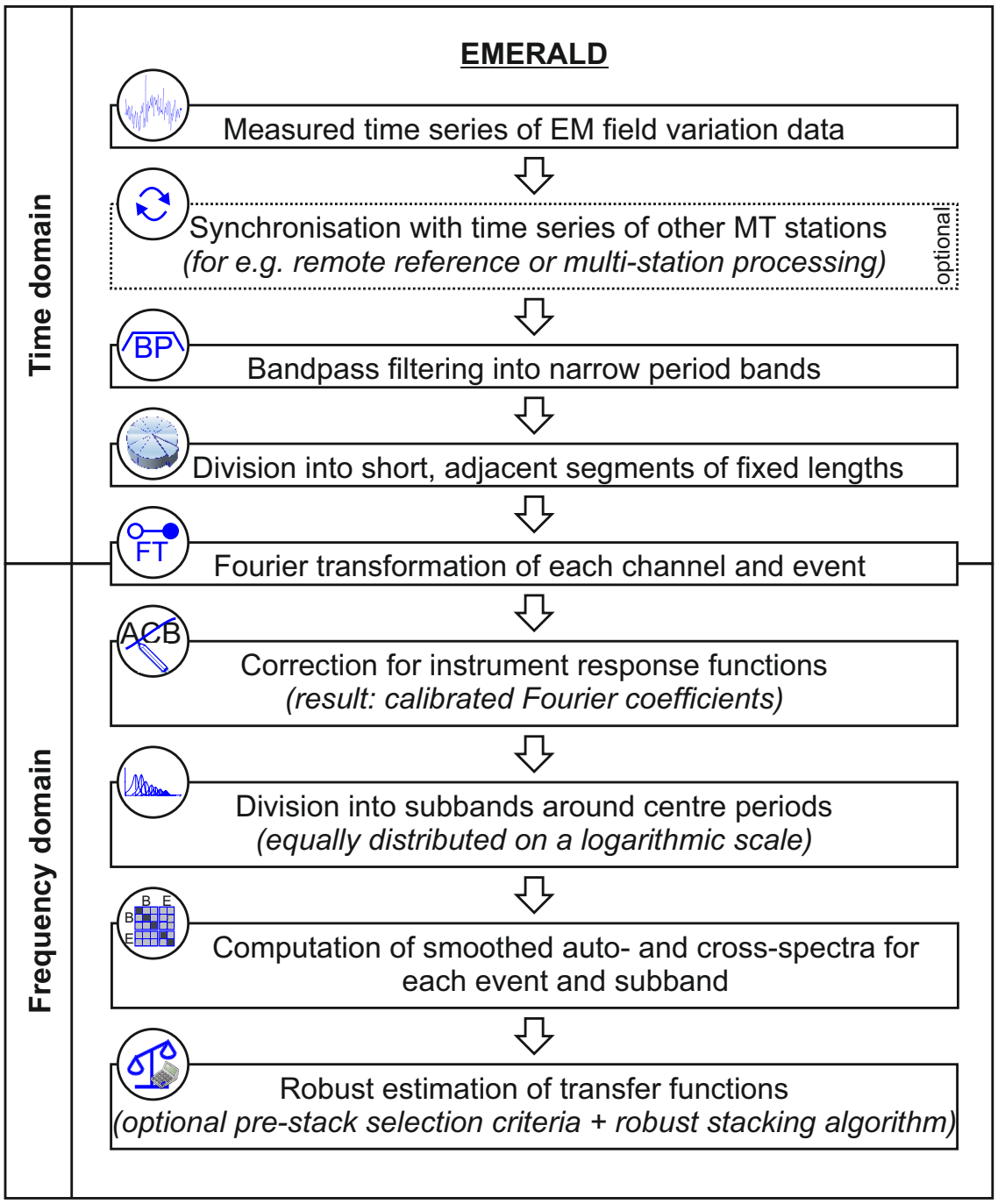

Figure A.1: Work flow of EMERALD processing (Ritter et al., 1998; Weckmann et al., 2005; Krings, 2007).

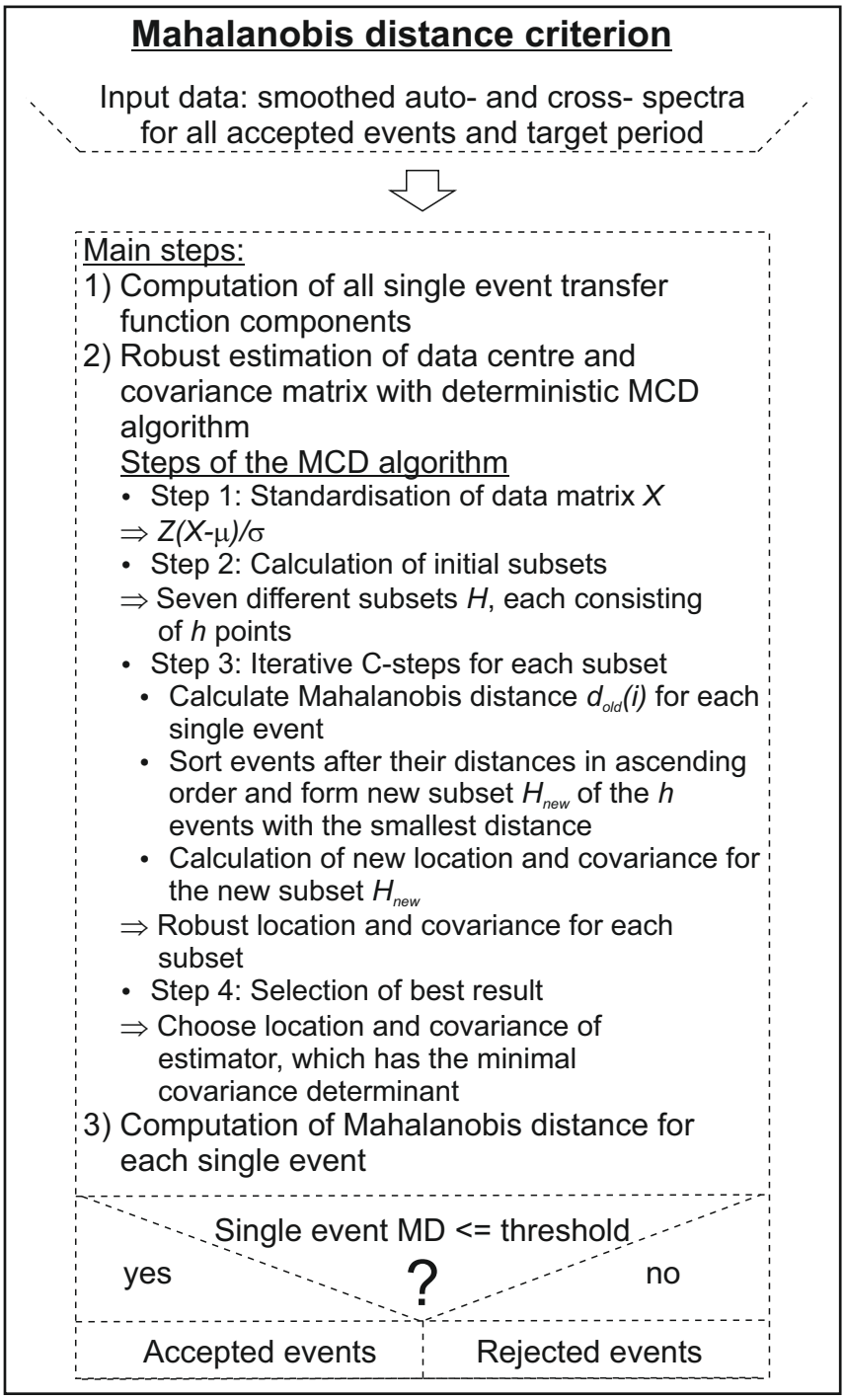

Figure A.2: Work flow of the Mahalanobis distance (MD) criterion.

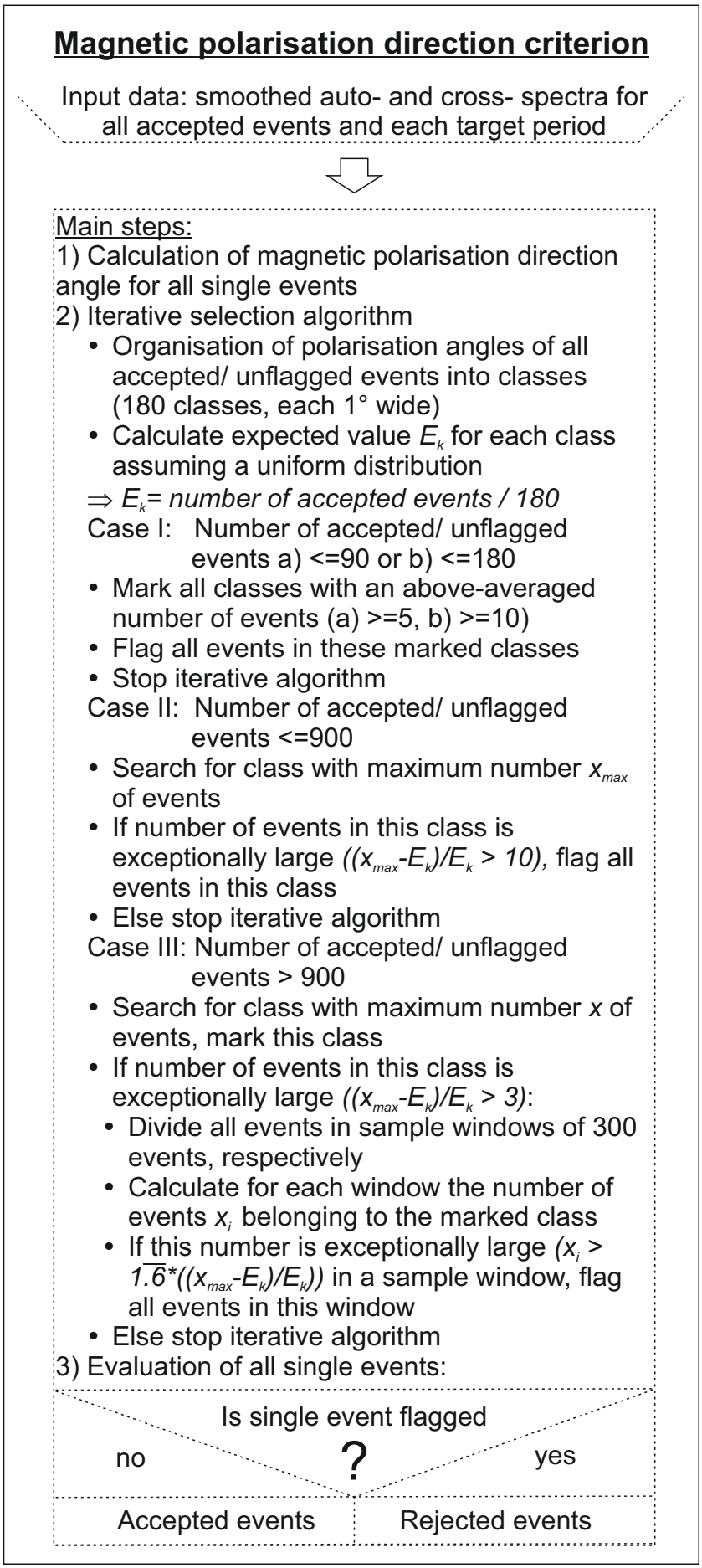

Figure A.3: Work flow of the magnetic polarisation direction (MPD) criterion.

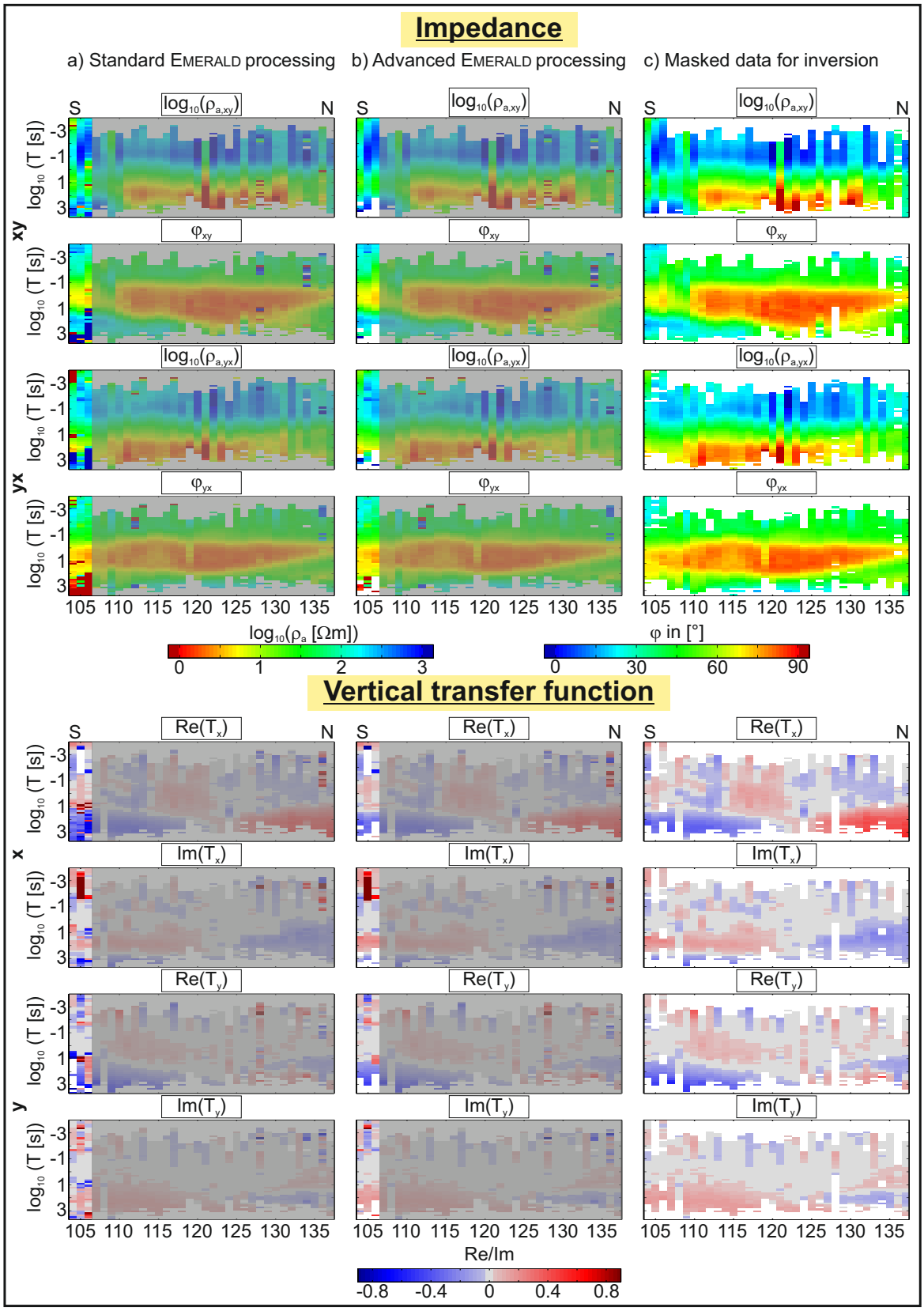

Figure A.4: Processing results of impedances (off-diagonal components) and vertical transfer functions for profile 1. A coherence threshold of 0.9 was applied for all processings. Columns from left to right: a) standard EMERALD processing, b) advanced EMERALD processing using MD and MPD criterion, c) masked data used for the inversion. Only the first three stations were processed in the framework of this thesis as the other stations were measured and evaluated by Weckmann et al. (2007a). White space represents not existing and masked data.

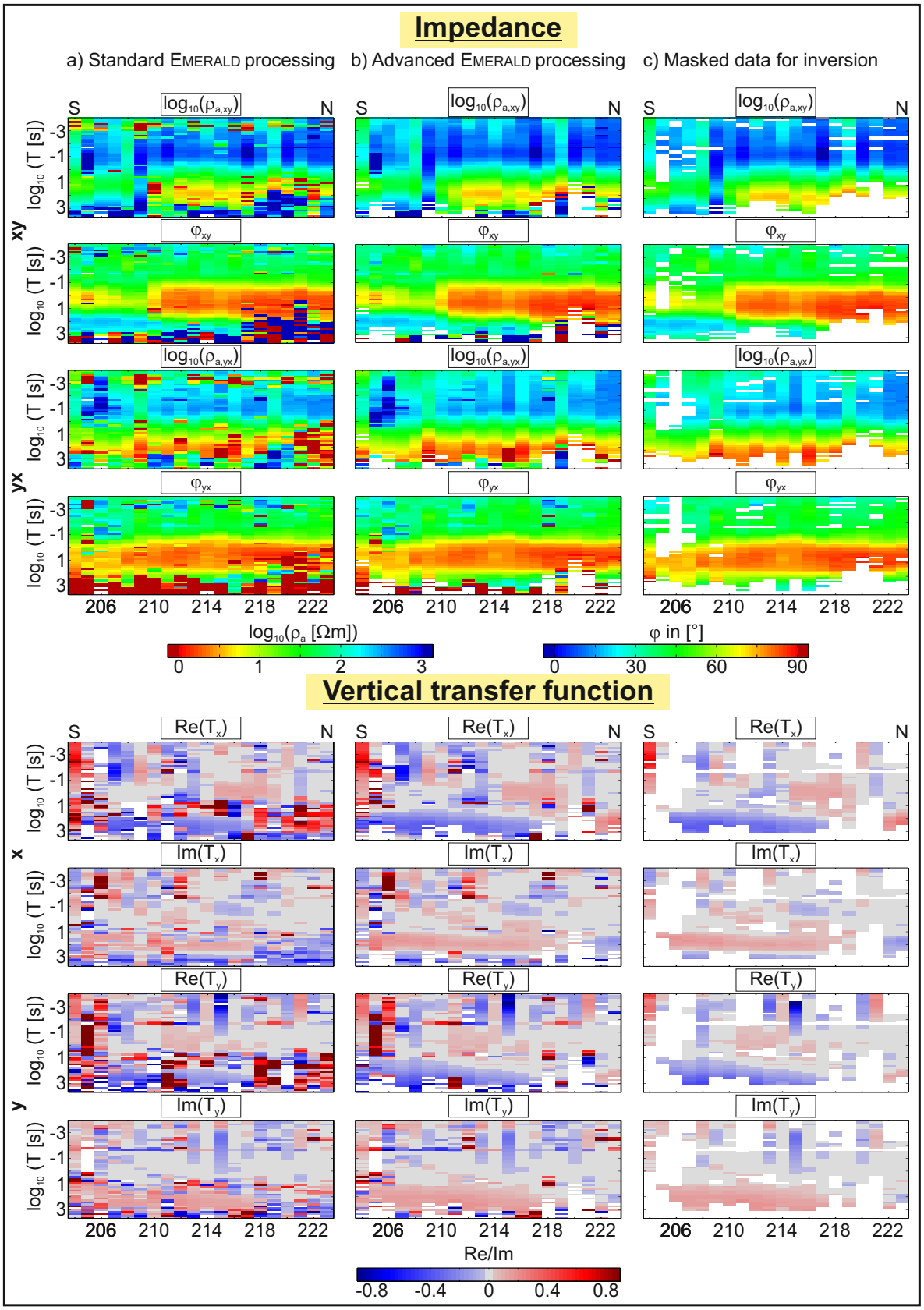

Figure A.5: Processing results of impedances (off-diagonal components) and vertical transfer functions for profile 2. A coherence threshold of 0.9 was applied for all processings. Columns from left to right: a) standard EMERALD processing, b) advanced EMERALD processing using MD and MPD criterion, c) masked data used for the inversion. White space represents not existing and masked data.

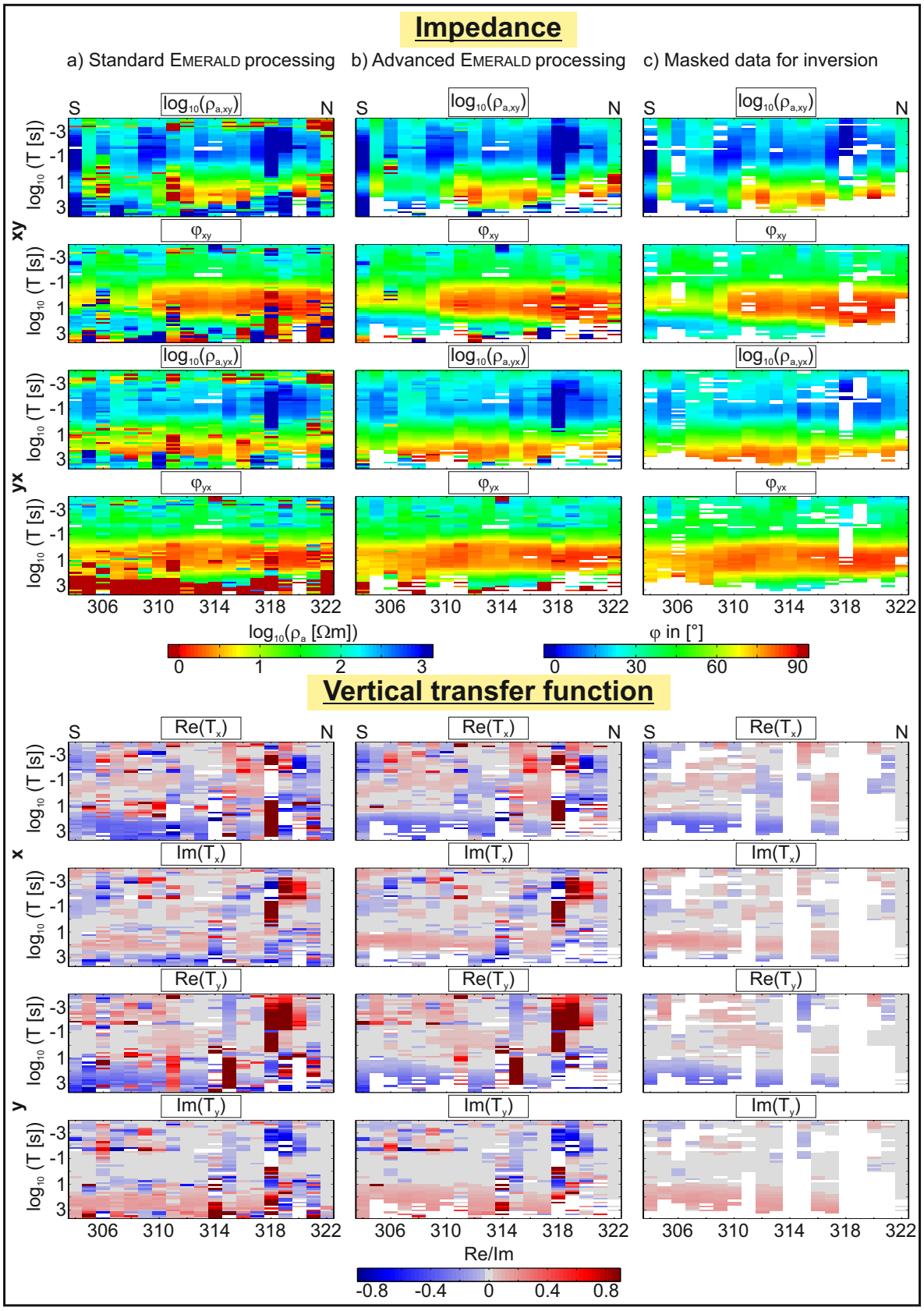

Figure A.6: Processing results of impedances (off-diagonal components) and vertical transfer functions for profile 3. A coherence threshold of 0.9 was applied for all processings. Columns from left to right: a) standard EMERALD processing, b) advanced EMERALD processing using MD and MPD criterion, c) masked data used for the inversion. White space represents not existing and masked data.

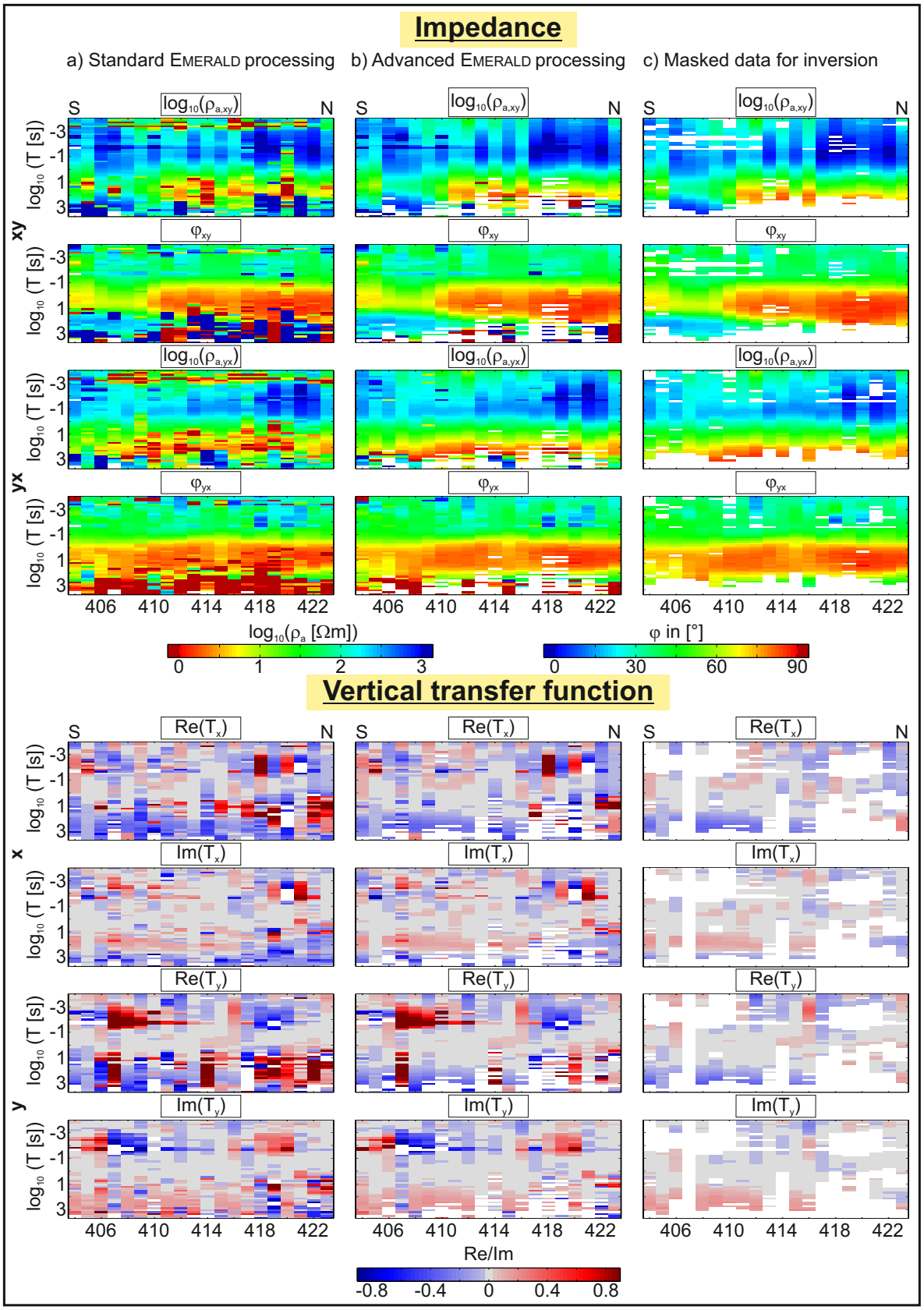

Figure A.7: Processing results of impedances (off-diagonal components) and vertical transfer functions for profile 4. A coherence threshold of 0.9 was applied for all processings. Columns from left to right: a) standard EMERALD processing, b) advanced EMERALD processing using MD and MPD criterion, c) masked data used for the inversion. White space represents not existing and masked data.

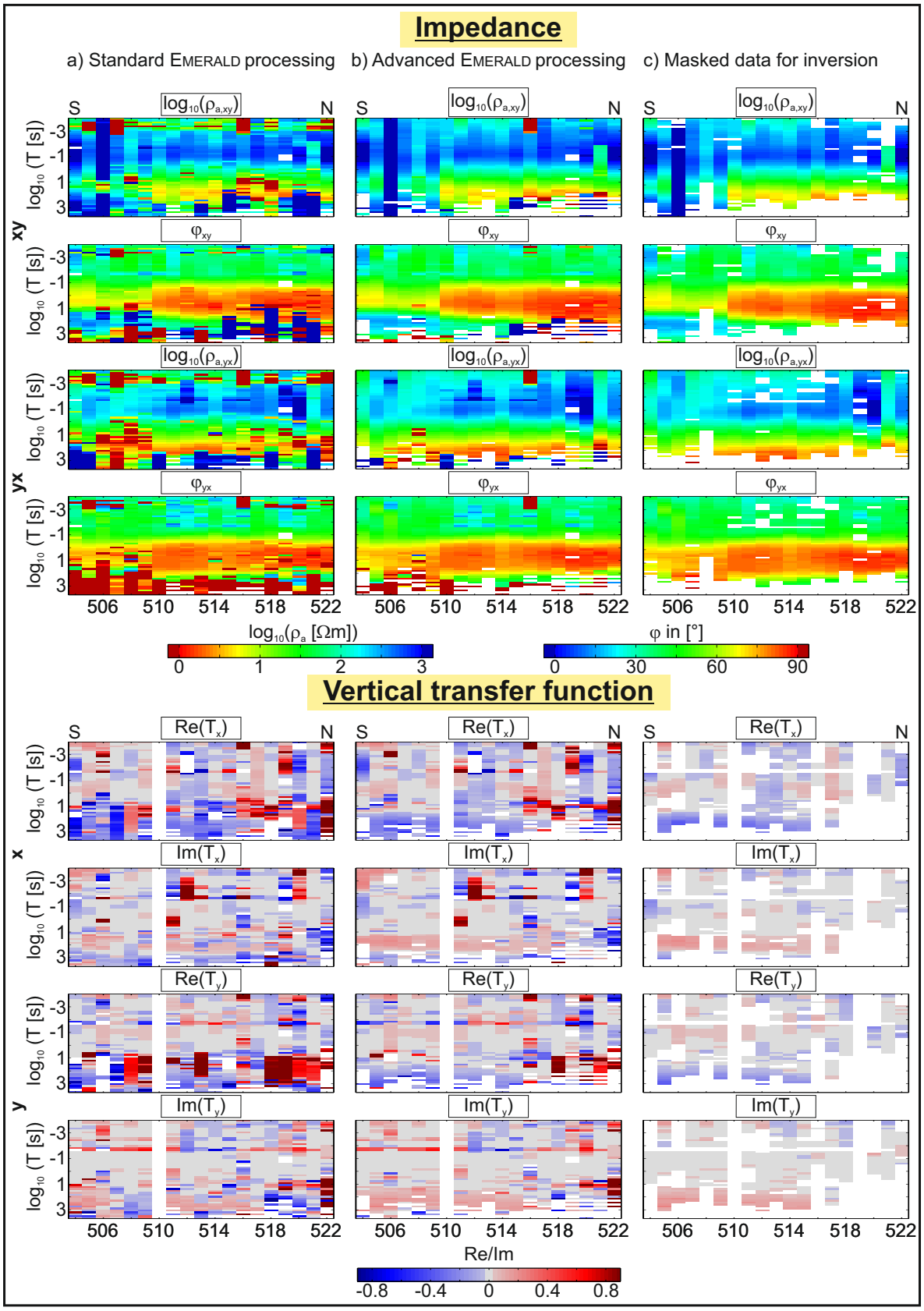

Figure A.8: Processing results of impedances (off-diagonal components) and vertical transfer functions for profile 5. A coherence threshold of 0.9 was applied for all processings. Columns from left to right: a) standard EMERALD processing, b) advanced EMERALD processing using MD and MPD criterion, c) masked data used for the inversion. White space represents not existing and masked data.

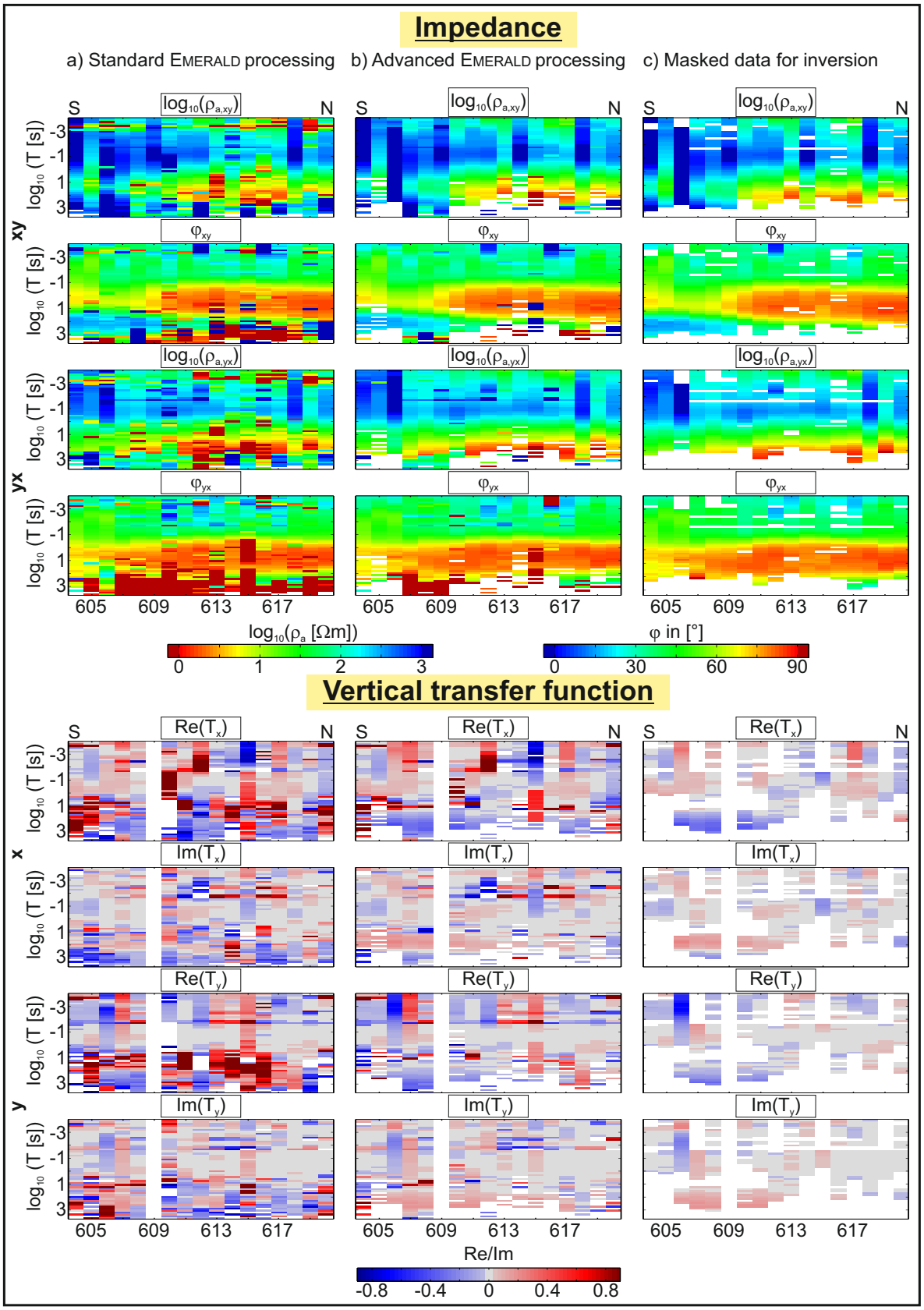

Figure A.9: Processing results of impedances (off-diagonal components) and vertical transfer functions for profile 6. A coherence threshold of 0.9 was applied for all processings. Columns from left to right: a) standard EMERALD processing, b) advanced EMERALD processing using MD and MPD criterion, c) masked data used for the inversion. White space represents not existing and masked data.

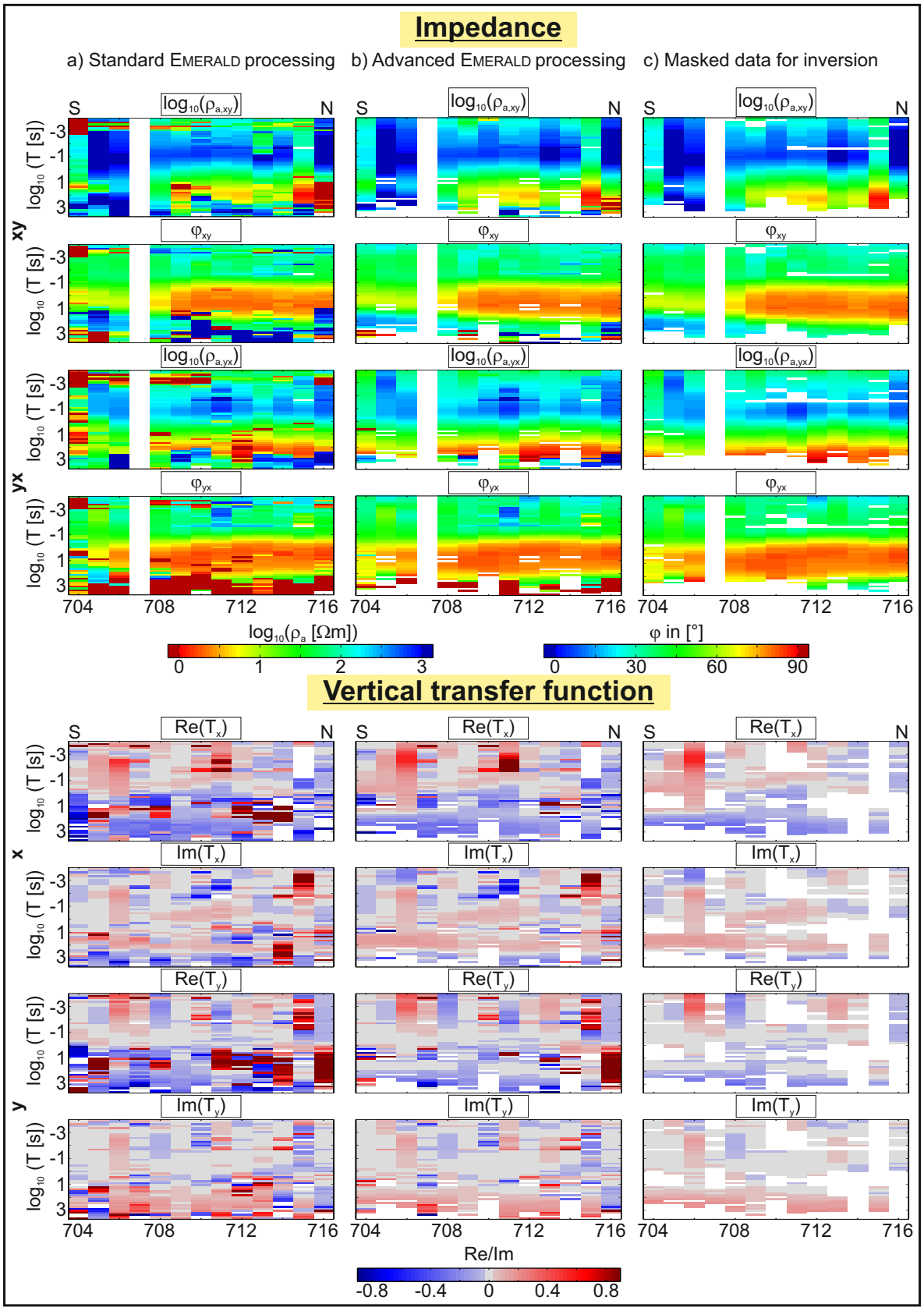

Figure A.10: Processing results of impedances (off-diagonal components) and vertical transfer functions for profile 7. A coherence threshold of 0.9 was applied for all processings. Columns from left to right: a) standard EMERALD processing, b) advanced EMERALD processing using MD and MPD criterion, c) masked data used for the inversion. White space represents not existing and masked data.

Bibliography

Adao, F., 2015. *The electrical resistivity of the Posidonia black shale*, PhD thesis, Freie Universität Berlin, Berlin.

Amidan, B. G., Ferryman, T. A., & Cooley, S. K., 2005. Data outlier detection using the Chebyshev theorem, in *Aerospace Conference, 2005 IEEE*, pp. 3814–3819, IEEE, Piscataway, N.J.

Basu, M. & Ho, T. K., 2006. *Data Complexity in Pattern Recognition*, Springer London.

Beattie, J., 1909. *Report of a magnetic survey of South Africa*, Cambridge University Press, London.

Branch, T., Ritter, O., Weckmann, U., Sachsenhofer, R. F., & Schilling, F., 2007. The Whitehill Formation a high conductivity marker horizon in the Karoo Basin, *South African Journal of Geology*, **110**(2-3), 465–476.

Brereton, R. G., 2015. The Mahalanobis distance and its relationship to principal component scores, *Journal of Chemometrics*, **29**(3), 143–145.

Cagniard, L., 1953. Basic theory of the Magneto-Telluric Method of geophysical prospecting, *Geophysics*, **18**, 605–635.

Campanyà, J., Ogaya, X., Jones, A. G., Rath, V., Vozar, J., & Meqbel, N., 2016. The advantages of complementing MT profiles in 3-D environments with geomagnetic transfer function and interstation horizontal magnetic transfer function data: Results from a synthetic case study, *Geophysical Journal International*, **207**(3), 1818–1836.

Catuneanu, O., Wopfner, H., Eriksson, P. G., Cairncross, B., Rubidge, B. S., Smith, R., & Hancox, P. J., 2005. The Karoo basins of south-central Africa, *Journal of African Earth Sciences*, **43**(1-3), 211–253.

Chave, A. D., 2014. Magnetotelluric data, stable distributions and impropriety: An existential combination, *Geophysical Journal International*, **198**(1), 622–636.

Chave, A. D. & Jones, A. J., 2012. *The magnetotelluric method*, Cambridge University Press.

Chave, A. D. & Thomson, D. J., 2004. Bounded influence magnetotelluric response function estimation, *Geophysical Journal International*, **157**(3), 988–1006.

Chave, A. D., Thomson, D. J., & Ander, M. E., 1987. On the robust estimation of power spectra,

coherences, and transfer functions, *Journal of Geophysical Research*, **92**(B1), 633.

Chen, X., 2007. A New Generalization of Chebyshev Inequality for Random Vectors, *arXiv:0707.0805*.

Chen, X., 2008. *Filterung von geophysikalischen Zeitreihen mit periodisch auftretenden multifrequenten Stoersignalen*, Diploma thesis, University of Potsdam, Potsdam.

Cole, D. J., 1992. Evolution and development of the Karoo Basin, in *Inversion tectonics of the Cape Fold Belt, Karoo and Cretaceous basins of southern Africa*, pp. 87–99, A.A. Balkema, Rotterdam.

de Beer, J. & Meyer, R., 1983. Geoelectrical and gravitational characteristics of the Namaqua-Natal Mobile Belt and its boundaries, *Spec. Publ. Geol. Soc. S. Afr.*, **10**, 91–100.

de Beer, J. & Meyer, R., 1984. Geophysical characteristics of the Namaqua-Natal Belt and its boundaries, South Africa, *Journal of Geodynamics*, **1**(3-5), 473–494.

de Maesschalck, R., Jouan-Rimbaud, D., & Massart, D. L., 2000. The Mahalanobis distance, *Chemometrics and Intelligent Laboratory Systems*, **50**(1), 1–18.

eds de Wit, M. J. & Ransome, I., 1992. *Inversion tectonics of the Cape Fold Belt, Karoo and Cretaceous basins of southern Africa*, A.A. Balkema, Rotterdam.

Donoho, D. & Huber, P., 1983. The notion of breakdown point, in *A Festschrift for Erich L. Lehmann*, Wadsworth statistics : probability series, eds Bickel, P. J. & Lehmann, E. L., Wadsworth, Belmont Calif.

Egbert, G. D., 1997. Robust multiple-station magnetotelluric data processing, *Geophysical Journal International*, **130**(2), 475–496.

Egbert, G. D. & Booker, J. R., 1986. Robust estimation of geomagnetic transfer functions, *Geophys. J. R. astr. Soc.*, pp. 173–194.

Egbert, G. D. & Kelbert, A., 2012. Computational recipes for electromagnetic inverse problems, *Geophysical Journal International*, **189**(1), 251–267.

ed. EIA, 2013. *Technically Recoverable Shale Oil and Shale Gas Resources: An Assessment of 137 Shale Formations in 41 Countries Outside the United States*, Washington.

Eilers, P. H. C. & Goeman, J. J., 2004. Enhancing scatterplots with smoothed densities, *Bioinformatics (Oxford, England)*, **20**(5), 623–628.

Falk, M., 1997. On Mad and Comedians, *Annals of the Institute of Statistical Mathematics*, **49**(4), 615–644.

Filzmoser, P., Garrett, R. G., & Reimann, C., 2005. Multivariate outlier detection in exploration geochemistry, *Computers & Geosciences*, **31**(5), 579–587.

Friebel, T., Stockmann, M., & Haber, R., 2010. Sensorueberwachung mit einer robusten zweidimensionalen Regelkarte, in *AALE 2010*, pp. 71–76.

Gamble, T. D., Goubau, W. M., & Clarke, J., 1979. Magnetotellurics with a remote magnetic reference, *Geophysics*, **44**(1), 53–68.

Geel, C., 2014. *Shale gas characteristics of Permian black shales in the Ecca Group, near Jansenville, Eastern Cape, South Africa*, Master thesis, Nelson Mandela Metropolitan University, Port Elisabeth.

Geel, C., Schulz, H.-M., Booth, P., deWit, M., & Horsfield, B., 2013. Shale Gas Characteristics of Permian Black Shales in South Africa: Results from Recent Drilling in the Ecca Group (Eastern Cape), *Energy Procedia*, **40**, 256–265.

Gnanadesikan, R. & Kettenring, J. R., 1972. Robust Estimates, Residuals, and Outlier Detection with Multiresponse Data, *Biometrics*, **28**(1), 81.

Goubau, W. M., Gamble, T. D., & Clarke, J., 1978. Magnetotelluric data analysis: Removal of bias, *Geophysics*, **43**(6), 1157–1166.

Greve, J., in preparation. *Magnetotelluric Modelling of the Eastern Karoo Basin Near Jansenville, South Africa: Imaging a Potential Shale Gas Bearing Horizon*, Master thesis, Nelson Mandela University, Port Elisabeth.

Hampel, F. R., 1986. *Robust statistics: The approach based on influence functions*, Wiley series in probability and mathematical statistics : Probability and mathematical statistics, Wiley, New York NY u.a.

Huber, P. J., 1981. *Robust statistics*, Wiley series in probability and mathematical statistics : Probability and mathematical statistics, Wiley, New York NY u.a.

Hubert, M. & Debruyne, M., 2010. Minimum covariance determinant, *Wiley Interdisciplinary Reviews: Computational Statistics*, **2**(1), 36–43.

Hubert, M., Rousseeuw, P. J., & van Aelst, S., 2008. High-Breakdown Robust Multivariate Methods, *Statistical Science*, **23**(1), 92–119.

Hubert, M., Rousseeuw, P. J., & Verdonck, T., 2012. A Deterministic Algorithm for Robust Location and Scatter, *Journal of Computational and Graphical Statistics*, **21**(3), 618–637.

Johnson, M., van Vuuren, C., Hegenberger, W., Key, R., & Show, U., 1996. Stratigraphy of the Karoo Supergroup in southern Africa: An overview, *Journal of African Earth Sciences*, **23**(1), 3–15.

Johnson, M., Vuuren, C., Visser, J., Cole, D., Wickens, H. d. V., Christie, A., Roberts, D., & Brandl, G., 2006. Sedimentary rocks of the Karoo Supergroup, in *The geology of South Africa*, ed. Johnson, M. R., Geological Soc. of South Africa [u.a.], Johannesburg.

Jones, A. G., Chave, A. D., Egbert, G., Auld, D., & Bahr, K., 1989. A comparison of techniques for magnetotelluric response function estimation, *Journal of Geophysical Research: Solid Earth*, **94**(B10), 14201–14213.

Junge, A., 1996. Characterization of and correction for cultural noise, *Surveys in Geophysics*, **17**(4), 361–391.

Kapinos, G., Weckmann, U., Jegen-Kulcsar, M., Meqbel, N., Neska, A., Katjiuongua, T. T., Hoelz, S., & Ritter, O., 2016. Electrical resistivity image of the South Atlantic continental margin derived from onshore and offshore magnetotelluric data, *Geophysical Research Letters*, **43**(1), 154–160.

Kappler, K. N., Morrison, H. F., & Egbert, G. D., 2010. Long-term monitoring of ULF electromagnetic fields at Parkfield, California, *Journal of Geophysical Research: Solid Earth*, **115**(B4).

Kaufmann, A. A. & Keller, G. V., 1981. *The magnetotelluric sounding method*, vol. 15 of **Methods in Geochemistry and Geophysics**, Elsevier.

Kelbert, A., Meqbel, N., Egbert, G. D., & Tandon, K., 2014. ModEM: A modular system for inversion of electromagnetic geophysical data, *Computers & Geosciences*, **66**, 40–53.

Kendall, M. & Buckland, W., 1957. *A dictionary of statistical terms*, Oliver and Boyd, London.

Korolevski, W., Ritter, O., Weckmann, U., Rybin, A., & Matiukov, V., 2014. Magnetotelluric Study of the Southern Pamir, Tajikistan, *AGU Fall Meeting Abstracts*.

Krings, T., 2007. *The influence of Robust Statistics, Remote Reference, and Horizontal Magnetic*

Transfer Functions on data processing in Magnetotellurics, Diploma thesis, University Muenster, Muenster.

Kuetter, S., 2015. *Magnetotelluric measurements across the southern Barberton Greenstone Belt, South Africa*, PhD thesis, University of Potsdam, Potsdam.

Larsen, J. C., 1989. Transfer functions: Smooth robust estimates by least-squares and remote reference methods, *Geophysical Journal International*, **99**(3), 645–663.

Lehmann, R., 2012. *Der Einfluss statistischer Ausreisser auf die Schaetzung der natuerlichen Variabilitaet in Daten zu Biota*, PhD thesis, RWTH Aachen University, Aachen.

Lock, B. E., 1978. The Cape Fold Belt of South Africa; tectonic control of sedimentation, *Proceedings of the Geologists' Association*, **89**(4), 263–281.

Lohninger, H., 2012. *Fundamentals of Statistics*, Epina e-Book Team.

Mahalanobis, P., 1936. On the generalized distance in statistics, *Proceedings of the National Institute of Science of India*.

ed. Malisa, V., 2010. *AALE 2010*.

Maronna, R. A. & Zamar, R. H., 2002. Robust Estimates of Location and Dispersion for High-Dimensional Datasets, *Technometrics*, **44**(4), 307–317.

Marshall, A. W. & Olkin, I., 1960. Multivariate Chebyshev Inequalities, *The Annals of Mathematical Statistics*, **31**(4), 1001–1014.

Meqbel, N., 2009. *The electrical conductivity structure of the Dead Sea Basin derived from 2D and 3D inversion of magnetotelluric data*, PhD thesis, Freie Universität Berlin, Berlin.

Meqbel, N., Egbert, G., Wannamaker, P., Kelbert, A., & Schultz, A., 2014. Deep electrical resistivity structure of the northwestern U.S. derived from 3-D inversion of USArray magnetotelluric data, *Earth and Planetary Science Letters*, **402**, 290–304.

Morrison, D. F., 1967. *Multivariate statistical methods*, McGraw-Hill.

Muñoz, G., Ritter, O., & Moeck, I., 2010. A target-oriented magnetotelluric inversion approach for characterizing the low enthalpy Groß Schönebeck geothermal reservoir, *Geophysical Journal International*, **183**(3), 1199–1215.

Newman, G., Gasperikova, E., Hoversten, M., & Wannamaker, P., 2008. Three-dimensional magnetotelluric characterization of the Coso geothermal field, *Geothermics*, **37**(4), 369–399.

Oettinger, G., Haak, V., & Larsen, J. C., 2001. Noise reduction in magnetotelluric time-series with a new signal-noise separation method and its application to a field experiment in the Saxonian Granulite Massif, *Geophysical Journal International*, **146**(3), 659–669.

Platz, A. & Weckmann, U., In revision. The Mahalanobis distance: A new measure to detect outliers in Magnetotelluric data, *Geophysical Journal International*.

Ritter, O., Junge, A., & Dawes, G., 1998. New equipment and processing for magnetotelluric remote reference observations, *Geophysical Journal International*, **132**(3), 535–548.

Rousseeuw, P., 1985. Multivariate Estimation with High Breakdown Point, in *Mathematical Statistics and Applications*, pp. 283–297, eds Grossmann, W., Pflug, G. C., Vincze, I., & Wertz, W., Springer Netherlands, Dordrecht.

Rousseeuw, P. J., 1984. Least Median of Squares Regression, *Journal of the American Statistical Association*, **79**(388), 871–880.

Rousseeuw, P. J. & Molenberghs, G., 1993. Transformation of non positive semidefinite correlation matrices, *Communications in Statistics - Theory and Methods*, **22**(4), 965–984.

Rousseeuw, P. J. & van Driessen, K., 1999. A Fast Algorithm for the Minimum Covariance Determinant Estimator, *Technometrics*, **41**(3), 212.

Sass, P., 2013. *Magnetotellurische Untersuchung der kontinentalen Kollisionszone im Pamir und Tian Shan, Zentralasien*, PhD thesis, Freie Universität Berlin, Berlin.

Schmitz, M., Orihuela, N. D., Klarica, S., Gil, E., Levander, A., Audemard, F. A., Mazuera, F., & Avila, J., 2013. Lithospheric scale model of Merida Andes, Venezuela (GIAME Project), *AGU Spring Meeting Abstracts*.

Schmucker, U., 1978. Auswertungsverfahren Göttingen, in *Haak, Homilius (Hg.) 1978 – Proceedings of the Colloquium on Electromagnetic Depth Sounding*.

Simpson, F. & Bahr, K., 2005. *Practical magnetotellurics*, Cambridge Univ. Press, Cambridge, 1st edn.

Sims, W. E., Bostick, F. X., & Smith, H. W., 1971. The Estimation of Magnetotelluric Impedance Tensor Elements from Measured Data, *Geophysics*, **36**(5), 938–942.

Siripunvaraporn, W., Egbert, G., Lenbury, Y., & Uyeshima, M., 2005a. Three-dimensional magnetotelluric inversion: Data-space method, *Physics of the Earth and Planetary Interiors*, **150**(1-3), 3–14.

Siripunvaraporn, W., Egbert, G., & Uyeshima, M., 2005b. Interpretation of two-dimensional magnetotelluric profile data with three-dimensional inversion: synthetic examples, *Geophysical Journal International*, **160**, 804–814.

Smirnov, M. Y., 2003. Magnetotelluric data processing with a robust statistical procedure having a high breakdown point, *Geophysical Journal International*, **152**(1), 1–7.

Srinivasaraghavan, J. & Allada, V., 2006. Application of mahalanobis distance as a lean assessment metric, *The International Journal of Advanced Manufacturing Technology*, **29**(11-12), 1159–1168.

Stankiewicz, J., Ryberg, T., Schulze, A., Lindeque, A., Weber, M. H., & de Wit, M. J., 2007. Initial results from wide-angle seismic refraction lines in the southern Cape, *South African Journal of Geology*, **110**(2-3), 407–418.

Stellato, B., van Parys, B. P. G., & Goulart, P. J., 2016. Multivariate Chebyshev Inequality with Estimated Mean and Variance, *The American Statistician*, pp. 1–13.

Svensen, H., Planke, S., Chevallier, L., Malthe-Sørenssen, A., Corfu, F., & Jamtveit, B., 2007. Hydrothermal venting of greenhouse gases triggering Early Jurassic global warming, *Earth and Planetary Science Letters*, **256**(3-4), 554–566.

Swift & Moore, C., 1967. *A magnetotelluric investigation of an electrical conductivity anomaly in the southwestern United States*, PhD thesis, Massachusetts Institute of Technology.

Szarka, L., 1988. Geophysical aspects of man-made electromagnetic noise in the earth—A review, *Surveys in Geophysics*, **9**(3-4), 287–318.

Thomas, R. J., von Veh, M. W., & McCourt, S., 1993. The tectonic evolution of southern Africa: An overview, *Journal of African Earth Sciences (and the Middle East)*, **16**(1-2), 5–24.

Tietze, K., 2012. *Investigating the electrical conductivity structure of the San Andreas fault system in the Parkfield-Cholame region, central California, with 3D magnetotelluric inversion*, PhD thesis,

Freie Universität Berlin, Berlin.

Tikhonov, A., 1950. On determining electrical characteristics of the deep layers of the earth's crust, *Doklady*, **73**, 295–297.

Tinker, J., de Wit, M., & Brown, R., 2008. Mesozoic exhumation of the southern Cape, South Africa, quantified using apatite fission track thermochronology, *Tectonophysics*, **455**(1-4), 77–93.

Travassos, J. M. & Beamish, D., 1988. Magnetotelluric data processing–a case study, *Geophysical Journal International*, **93**(2), 377–391.

van Zijl, J., 1978. The relationship between the deep electrical resistivity structure and tectonic provinces in Southern Africa - Part 1. Results obtained by Schlumberger soundings, *South African Journal of Geology*, **81**(2), 129–142.

van Zijl, J., 2006. A review of the resistivity structure of the Karoo Supergroup, South Africa, with emphasis on the dolerites: A study in anisotropy, *South African Journal of Geology*, **109**(3), 315–328.

Verboven, S. & Hubert, M., 2010. MATLAB library LIBRA, *Wiley Interdisciplinary Reviews: Computational Statistics*, **2**(4), 509–515.

Weckmann, U., 2015. *Die elektrische Leitfaehigkeit von fossilen Stoerungszonen und Mobile Belts*, Habilitationsschrift, University of Potsdam, Potsdam.

Weckmann, U., Magunia, A., & Ritter, O., 2005. Effective noise separation for magnetotelluric single site data processing using a frequency domain selection scheme, *Geophysical Journal International*, **161**(3), 635–652.

Weckmann, U., Jung, A., Branch, T., & Ritter, O., 2007a. Comparison of electrical conductivity structures and 2D magnetic modelling along two profiles crossing the Beattie Magnetic Anomaly, South Africa, *South African Journal of Geology*, **110**(2-3), 449–464.

Weckmann, U., Ritter, O., Jung, A., Branch, T., & de Wit, M., 2007b. Magnetotelluric measurements across the Beattie magnetic anomaly and the Southern Cape Conductive Belt, South Africa, *Journal of Geophysical Research*, **112**(B5).

Weckmann, U., Ritter, O., Chen, X., Tietze, K., & de Wit, M., 2012. Magnetotelluric image linked to surface geology across the Cape Fold Belt, South Africa, *Terra Nova*, **24**(3), 207–212.

Wiese, H., 1962. Geomagnetische Tiefentellurik Teil II: Die Streichrichtung der Untergrundstrukturen des elektrischen Widerstandes, erschlossen aus geomagnetischen Variationen, *Geofis. Pura e Appl.*, **52**, 83–103.

Xiao, Q., Cai, X., Xu, X., Liang, G., & Zhang, B., 2010. Application of the 3D magnetotelluric inversion code in a geologically complex area, *Geophysical Prospecting*, **58**(6), 1177–1192.